А. Г. КУРОШ

КУРС ВЫСШЕЙ АЛГЕБРЫ

ИЗДАНИЕ ДЕВЯТОЕ

Допущено Министерством высшего и среднего специального образования РСФСР в качестве учебника для университетов



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1968

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ К ШЕСТОМУ ИЗДАНИЮ

Первое издание этой книги вышло в 1946 г., а затем она переиздавалась в 1950, 1952, 1955 и 1956 гг. Перед вторым и четвертым изданиями книга подвергалась значительной переработке, имевшей целью отразить опыт алгебраического преподавания в Московском университете. При подготовке к настоящему шестому изданию книга подверглась еще более серьезной переработке, столь серьезной, что с достаточными основаниями ее можно было бы считать новой книгой, а не шестым изданием старой книги.

Эта переработка определялась двумя задачами. Прежде всего, неоднократно высказывались пожелания о расширении книги для того, чтобы она обеспечивала весь обязательный университетский курс высшей алгебры, а не только его первые два семестра, как это было до сих пор. С этой целью в книгу включено несколько новых глав. Одна из них посвящена основам теории групп, а остальные относятся к линейной алгебре — теория линейных пространств, теория евклидовых пространств, теория λ-матриц и жордановой нормальной формы матрицы.

Конечно, в советской алгебраической литературе в настоящее время имеется ряд хороших книг по линейной алгебре, различных по объему, содержанию, характеру изложения. Настоящая книга, даже после столь значительного добавления к ней материала, относящегося к линейной алгебре, не может претендовать на замену какой-либо из этих книг. Тем не менее бесспорно, что студентам будет удобно иметь весь обязательный материал собранным в одном учебнике и изложенным единым стилем.

С другой стороны, расположение глав, принятое в предшествующих изданиях книги, уже давно не соответствует действующему в Московском университете фактическому порядку изложения материала — этот последний в большой мере определяется необходимостью к определенному сроку выполнять определенные заказы курсов аналитической геометрии и математического анализа. Больше того, три года тому назад в Московском университете была введена новая программа курса высшей алгебры. За эти годы она успешно прошла испытания и поэтому казалось целесообразным перестроить книгу, расположив в ней материал в точном соответствии с указанной программой. Появление учебника, соответствующего этой

программе, облегчит, вероятно, ее введение и в других университетах страны.

Укажем распределение материала по семестрам: 1-й семестр — главы 1—5, 2-й семестр — главы 6—9, 3-й семестр — главы 10, 11, 13 и 14. Следует отметить, что студенты-механики Московского университета изучают высшую алгебру лишь в объеме первых двух семестров.

Несомненно, что эта перестройка книги не затруднит, а, быть может, даже и облегчит ее использование в педагогических институтах.

Предшествующие переработки книги удавалось выполнить без всякого увеличения ее объема. На этот раз сделать это было, конечно, невозможно. Желание в какой-то мере сократить объем книги заставило исключить из нее некоторый материал, в частности параграфы, посвященные теореме Гурвица, теории алгебр и теореме Фробениуса. Тем не менее не казалось разумным ограничиться изложением в книге лишь того материала, который входит сейчас в обязательную программу, т. е. превратить эту книгу в простой конспект лекций. Сохраненный в книге необязательный материал — параграфы, целиком к нему относящиеся, отмечены звездочкой,— как правило, таков, что в свое время он входил в обязательную программу курса высшей алгебры, в некоторых университетах или педагогических институтах входит в программу и сейчас и во всяком случае был бы включен в программу, если бы курс высшей алгебры располагал бо́льшим числом часов.

При переработке книги были изменены также некоторые детали, на которых мы не будем, однако, сейчас останавливаться.

А. Курош

Москва, декабрь 1958 г.

ВВЕДЕНИЕ

Математическое образование студента-математика начинается с изучения трех основных дисциплин, а именно математического анализа, аналитической геометрии и высшей алгебры. Эти дисциплины имеют ряд точек соприкосновения, а местами и перекрытий, и вместе составляют фундамент, на котором строится все здание современной математической науки.

Высшая алгебра, изложению которой посвящена настоящая книга, представляет собой далеко идущее, но вполне естественное обобщение основного содержания школьного курса элементарной алгебры. Центральным в школьном курсе алгебры является, бесспорно, вопрос о решении уравнений. Как читатель помнит, изучение уравнений начинается с очень простого случая одного уравнения первой степени с одним неизвестным, а затем развивается в двух направлениях. С одной стороны, рассматриваются системы двух и трех уравнений первой степени с двумя и, соответственно, тремя неизвестными; с другой стороны, изучается одно квадратное уравнение с одним неизвестным, а также некоторые частные типы уравнений более высокой степени, легко сводящиеся к квадратным (биквадратные уравнения, например).

Эти оба направления получают дальнейшее развитие в курсе высшей алгебры, определяя ее разбиение на два больших отдела. Один из них, а именно основы линейной алгебры, имеет исходной задачей изучение произвольных систем уравнений первой степени или, как говорят, линейных уравнений. Для решения таких систем в том случае, когда число уравнений равно числу неизвестных, разрабатывается аппарат теории определителей. Этого аппарата уже недостаточно, однако, для изучения таких систем линейных уравнений, у которых число уравнений не равно числу неизвестных, — случай, непривычный с точки зрения элементарной алгебры, но очень важный для приложений. Оказалось необходимым, в частности, разрабатывать теорию матриц, т. е. систем чисел, расположенных в квадратные или прямоугольные таблицы из нескольких строк и столбцов. Эта теория оказалась очень глубокой и нашла приложения далеко за пределами теории систем линейных уравнений. С другой стороны, изучение систем линейных уравнений потребовало введения и изучения многомерных (так называемых векторных или линейных)

пространств. У людей, далеких от математики, с многомерным (в первую очередь с четырехмерным) пространством связываются туманные и часто ошибочные представления; в действительности же это понятие является чисто математическим, даже в основном алгебраическим, и служит важным орудием во многих математических исследованиях, а также в физике и механике.

Вторая половина курса высшей алгебры, называемая алгеброй многочленов, посвящена изучению одного уравнения от одного неизвестного, но уже произвольной степени. Учитывая существование формулы для решения квадратных уравнений, естественно было искать аналогичные формулы для уравнений более высоких степеней. Исторически этот отдел алгебры так и развивался, причем формулы для решения уравнений третьей и четвертой степени были найдены еще в XVI веке. После этого начались безуспешные поиски формул, которые выражали бы корни уравнений пятой и более высоких степеней через коэффициенты этих уравнений при помощи радикалов, быть может и очень многоэтажных. Эти поиски продолжались до начала XIX века, когда было, наконец, доказано, что такие формулы не могут быть найдены и что для всех степеней, начиная с пятой, существуют даже конкретные примеры уравнений с целочисленными коэффициентами, корни которых не могут быть записаны при помощи радикалов.

Отсутствие формул для решения уравнений высоких степеней не следует считать очень печальным обстоятельством — даже в случае уравнений третьей и четвертой степени, где такие формулы существуют, они очень громоздки и практически почти бесполезны. С другой стороны, коэффициенты тех уравнений, которые приходится решать физикам или инженерам, являются обычно величинами, полученными в результате измерений, т. е. известны лишь приближенно, а поэтому и корни нужно знать лишь приближенно, с заданной точностью. Это привело к разработке различных методов приближенного решения уравнений, лишь простейшие из которых излагаются в курсе высшей алгебры.

Центральным в алгебре многочленов оказывается, однако, не вопрос о практическом разыскании корней уравнений, а вопрос об их существовании. Известно, что существуют даже квадратные уравнения с действительными коэффициентами, не имеющие действительных корней. Пополняя запас чисел до совокупности всех комплексных чисел, мы обнаруживаем, что квадратные уравнения уже корнями обладают и что это же справедливо и для уравнений третьей и четвертой степени, как вытекает из существования формул для их решения. Не найдется ли, однако, такое уравнение пятой или более высокой степени, которое не имеет ни одного корня даже среди комплексных чисел, и не придется ли для разыскания корней подобных уравнений переходить от комплексных чисел к более широкому запасу чисел? Ответ на этот вопрос дает важная теорема, утвер-

ждающая, что всякое уравнение с любыми числовыми коэффициентами, не только действительными, но и комплексными, имеет комплексные (быть может, в частности, действительные) корни, причем корней этих, вообще говоря, даже столько, какова степень уравнения.

Таков краткий обзор основного содержания курса высшей алгебры. Следует подчеркнуть, что высшая алгебра является лишь началом большой алгебраической науки, очень разветвленной, богатой содержанием и постоянно развивающейся. Попытаемся дать обзор, еще более беглый, тех ветвей алгебры, которые в основном лежат за пределами курса высшей алгебры.

Линейная алгебра, являющаяся большой наукой, посвященной в основном теории матриц и связанной с нею теории линейных преобразований векторных пространств, включает в себя также теорию форм, теорию инвариантов и тензорную алгебру, играющую важную роль в дифференциальной геометрии. Теория векторных пространств получает дальнейшее развитие вне алгебры, в функциональном анализе (бесконечномерные пространства). По разнообразию и значительности приложений как в математике, так и в механике, физике и технических науках линейная алгебра остается пока первой среди многочисленных ветвей алгебры.

Алгебра многочленов, развивавшаяся на протяжении многих десятилетий как наука об одном уравнении произвольной степени от одного неизвестного, теперь уже в основном закончена. Дальнейшее развитие она частично получила в некоторых разделах теории функций комплексного переменного, в основном же переросла в теорию полей, о которой скажем ниже. Что же касается очень трудного вопроса о системах уравнений от нескольких неизвестных, но не линейных, а произвольных степеней, — этот вопрос, объединяющий оба направления, разрабатываемые в курсе высшей алгебры, в самом этом курсе почти не затрагивается, — то он по существу относится к особой ветви математики, называемой алгебраической геометрией.

Исчерпывающее решение вопроса об условиях, при которых уравнение может быть решено в радикалах, было дано французским математиком Галуа (1811 — 1832). Его исследования указали новые направления в развитии алгебры, что привело уже в XX веке, после работ немецкой женщины-алгебраиста Э. Нётер (1882 — 1935), к оформлению новой точки зрения на задачи алгебраической науки. Сейчас бесспорно, что вовсе не изучение уравнений является центральной задачей алгебры. Истинным объектом алгебраического исследования следует считать алгебраические операции, подобные сложению или умножению чисел, но производимые, возможно, не над числами.

Уже школьнику приходится встречаться в курсе физики с операцией сложения сил. Математические дисциплины, изучаемые на первых курсах университетов и педагогических институтов, приносят многочисленные примеры алгебраических операций — сложение и умножение

матриц, функций, операции над преобразованиями пространства, над векторами и т. д. Эти операции обычно похожи на операции над числами и носят те же названия, но иногда некоторые свойства, привычные в случае чисел, оказываются утерянными. Так, очень часто и в очень важных случаях операции оказываются некоммутативными (произведение зависит от порядка сомножителей), а иногда и неассоциативными (произведение трех множителей зависит от расстановки скобок).

Наиболее систематическому изучению подвергаются немногие, наиболее важные типы алгебраических систем, т. е. множеств, составленных из элементов какой-либо природы, для которых определены некоторые алгебраические операции. Таковы, в частности, поля. Это будут алгебраические системы, в которых, подобно системе действительных и системе комплексных чисел, определены операции сложения и умножения, обе коммутативные и ассоциативные, связанные законом дистрибутивности (т. е. справедливо обычное правило раскрытия скобок) и обладающие обратными операциями — вычитанием и делением. Теория полей оказалась естественной областью для дальнейшего развития теории уравнений, а ее основные ветви — теория полей алгебраических чисел и теория полей алгебраических функций — связали ее соответственно с теорией чисел и теорией функций комплексного переменного. Курс высшей алгебры включает в себя элементарное введение в теорию полей, а некоторые разделы курса — многочлены от нескольких неизвестных, нормальная форма матрицы — излагаются сразу для случая произвольного основного поля.

Более широким, чем понятие поля, является понятие кольца. В отличие от случая поля, здесь уже не требуется выполнимости деления и, кроме того, умножение может быть некоммутативным и даже неассоциативным. Простейшими примерами колец служат совокупность всех целых чисел (включая и отрицательные), система многочленов от одного неизвестного и система действительных функций действительного переменного. Теория колец включает в себя такие старые ветви алгебры, как теория гиперкомплексных систем и теория идеалов, она связана с рядом математических наук, в частности с функциональным анализом, и уже нашла некоторые выходы в физику. Курс высшей алгебры, по существу, содержит лишь определение понятия кольца.

Еще бóльшую область применений имеет теория групп. Группой называется алгебраическая система с одной основной операцией, причем эта операция должна быть ассоциативной, хотя не обязательно коммутативной, и должна обладать обратной операцией — делением, если основная операция названа умножением. Такова, например, совокупность целых чисел, рассматриваемая относительно операции сложения, а также совокупность положительных действительных чисел, рассматриваемая с операцией умножения. Группы играли большую роль уже в теории Галуа, в вопросе о разрешимости уравнений в радикалах,

сейчас же они являются важным орудием в теории полей, во многих разделах геометрии, в топологии, а также и вне математики — в кристаллографии, в теоретической физике. Вообще, по широте области приложений теория групп занимает среди всех ветвей алгебры следующее после линейной алгебры место. Наш курс включает в себя главу, посвященную основам теории групп.

В самые последние десятилетия возникла и далеко развилась новая область алгебры — теория структур. Структурой называется алгебраическая система с двумя операциями — сложением и умножением. Эти операции должны быть коммутативными и ассоциативными, а также удовлетворять следующим требованиям: и сумма, и произведение элемента с самим собою должны равняться самому этому элементу; если сумма двух элементов равна одному из них, то произведение равно другому, и обратно. Примером структуры служит система натуральных чисел, рассматриваемая относительно операций взятия общего наименьшего кратного и общего наибольшего делителя. Теория структур имеет интересные связи с теорией групп и теорией колец, а также с теорией множеств; одна старая ветвь геометрии, а именно проективная геометрия, оказалась, по существу, частью теории структур; можно отметить также один выход теории структур в теорию электрических сетей.

Известный параллелизм, существующий между некоторыми частями теории групп, теории колец и теории структур, привел к возникновению общей теории алгебраических систем (или универсальных алгебр). Эта теория сделала пока лишь самые первые шаги, но контуры ее уже вырисовываются, а обнаружившиеся здесь связи с математической логикой позволяют рассчитывать на серьезное дальнейшее развитие.

Конечно, в изложенную выше схему далеко не укладывается все многообразное содержание алгебраической науки. Существует, в частности, ряд отделов алгебры, пограничных с другими разделами математики. Такова топологическая алгебра, изучающая алгебраические системы, в которых операции непрерывны относительно некоторой сходимости, определенной для элементов этих систем; примером служит система действительных чисел. К топологической алгебре близка теория непрерывных (или лиевых) групп, имеющая многочисленные приложения в различных вопросах геометрии, в теоретической физике, в гидродинамике. Впрочем, теория лиевых групп отличается таким переплетением алгебраических, топологических, геометрических и теоретико-функциональных методов, что было бы правильным считать ее особой ветвью математики. Существует, далее, теория упорядоченных алгебраических систем, возникшая в связи с исследованиями по основаниям геометрии и нашедшая приложения в функциональном анализе. Начинает развиваться, наконец, дифференциальная алгебра, устанавливающая новые связи между алгеброй и теорией дифференциальных уравнений.

Само собой разумеется, что то блестящее развитие алгебраической науки, которое привело к ее сегодняшнему состоянию, не было случайным — оно явилось частью общего развития математики и в значительной мере вызывалось необходимостью ответить на запросы, предъявляемые к алгебре со стороны других математических наук. С другой стороны, развитие алгебры само оказывало и оказывает очень большое влияние на развитие смежных ветвей науки, особенно усилившееся благодаря тому расширению области приложений, которое характерно для современной алгебры, и поэтому иногда говорят даже о происходящей сейчас «алгебраизации» математики.

Обзор алгебры, данный нами выше, не только является очень беглым, но и не дает представления об истории развития этой науки. Мы закончим поэтому наше введение очень кратким обзором истории алгебры.

Некоторыми вопросами алгебры, в частности решением простейших уравнений, занимались еще вавилонские, а затем древнегреческие математики. Вершиной алгебраических исследований этого периода являются сочинения греческого (александрийского) математика Диофанта (III век н. э.). В дальнейшем эти исследования развивались индийскими математиками — Ариабхата (VI век), Брамагупта (VII век), Бхаскара (XII век). Очень рано началась разработка вопросов алгебры в Китае — Чжан Цан (II век до н. э.), Цзин Чоу-чан (I век н. э.). Весьма крупным китайским алгебраистом был Цинь Цзю-шао (XIII век).

Большой вклад в развитие алгебры внесли математики средневекового востока, писавшие на арабском языке, в особенности уроженцы Средней Азии узбекский ученый Мухаммед Аль-Хорезми (IX век) и таджикский математик и поэт Омар Хайям (1040—1123). В частности, само слово «алгебра» возникло в связи с заглавием книги Аль-Хорезми «Аль-джебр аль-мукабала».

Упомянутые выше исследования вавилонских, греческих, индийских, китайских и среднеазиатских алгебраистов относились к тем вопросам алгебры, которые входят ныне в программу курса элементарной алгебры, и лишь иногда касались уравнений третьей степени. В этом же круге вопросов оставались в основном и исследования средневековых западноевропейских алгебраистов и алгебраистов эпохи Возрождения; назовем итальянского математика Леонардо Пизанского (Фибоначчи) (XII век) и создателя современной алгебраической символики француза Вьета (1540—1603). Впрочем, выше уже отмечалось, что в XVI веке были найдены методы решения уравнений третьей и четвертой степени; здесь должны быть названы имена итальянцев Ферро (1465—1526), Тарталья (1500—1557), Кардано (1501—1576) и Феррари (1522—1565).

В XVII и XVIII веках происходила интенсивная разработка общей теории уравнений (т. е. алгебры многочленов), в которой принимали участие крупнейшие ученые того времени — француз Декарт (1596—1650), англичанин Ньютон (1643—1727), французы Даламбер

(1717—1783) и Лагранж (1736—1813). В XVIII веке началось также построение теории определителей — швейцарский математик Крамер (1704—1752), французский ученый Лаплас (1749—1827). На рубеже XVIII и XIX веков немецкий математик Гаусс (1777—1855) доказал упоминавшуюся выше основную теорему о существовании корней уравнений с числовыми коэффициентами.

Первая треть XIX века ознаменована в истории алгебры решением проблемы о разрешимости уравнений в радикалах. Доказательство невозможности найти формулы для решения уравнений, степень которых больше или равна пяти, было получено итальянским математиком Руффини (1765—1822) и в более строгой форме норвежским ученым Абелем (1802—1829). Как уже отмечалось выше, исчерпывающее решение вопроса об условиях, при которых уравнение допускает решение в радикалах, принадлежит Галуа.

Теория Галуа явилась толчком для широкого развития алгебры в середине и второй половине XIX века, в том числе и новых ее направлений. Так, появились теория полей алгебраических чисел и полей алгебраических функций и связанная с ней теория идеалов. Здесь нужно назвать немецких математиков Куммера (1810—1893), Кронекера (1823—1891) и Дедекинда (1831—1916) и русских математиков Е. И. Золотарева (1847—1878) и Г. Ф. Вороного (1868—1908). Большое развитие получила теория конечных групп, идущая еще от Лагранжа и Галуа; здесь работали французы Коши (1789—1857) и Жордан (1838—1922), норвежский математик Силов (1832—1918), немецкие алгебраисты Фробениус (1849—1918) и Гёльдер (1859—1937). Начало теории непрерывных групп положили исследования норвежского математика С. Ли (1842—1899).

Работами английского ученого Гамильтона (1805—1865) и немецкого математика Грасмана (1809—1877) началась теория гиперкомплексных систем или, как теперь говорят, теория алгебр. Большую роль в дальнейшем развитии этой ветви алгебры играли относящиеся к концу века работы русского математика Ф. Э. Молина (1861—1941).

Линейная алгебра достигла в XIX веке большого расцвета, прежде всего благодаря работам английских математиков Сильвестра (1814—1897) и Кэли (1821—1895). Продолжалась разработка и алгебры многочленов; мы отметим лишь метод приближенного решения уравнений, найденный русским геометром Н. И. Лобачевским (1792—1856), и работы немецкого математика Гурвица (1859—1919). Во второй половине века начала создаваться алгебраическая геометрия, в частности в работах немецкого математика М. Нётера (1844—1922).

В XX веке алгебраические исследования приобрели очень большую широту и алгебра, как мы уже знаем, заняла в математике весьма почетное место. В этот период возникают многие новые разделы алгебры, в том числе общая теория полей (десятые годы), теория колец и общая теория групп (двадцатые годы), топологическая

алгебра и теория структур (тридцатые годы); в сороковых и пятидесятых годах появились теория полугрупп и теория квазигрупп, теория универсальных алгебр, гомологическая алгебра, теория категорий. Во всех частях алгебры работают крупные ученые, внесшие серьезный вклад в науку, в ряде стран возникают большие алгебраические школы. Это относится, в частности, к Советскому Союзу.

Из числа русских дореволюционных алгебраистов, помимо названных выше, следует указать также С. О. Шатуновского (1859—1929) и Д. А. Граве (1863—1939). Однако настоящий расцвет алгебраических исследований в нашей стране начинается лишь после Великой Октябрьской революции. Эти исследования захватывают почти все разделы современной алгебраической науки, причем в некоторых из них работы советских алгебраистов играют руководящую роль. Мы назовем лишь два имени — Н. Г. Чеботарева (1894—1947), работавшего в теории полей и теории лиевых групп, и О. Ю. Шмидта (1891—1956), известного полярника и в то же время крупного алгебраиста, создателя советской теоретико-групповой школы.

Заканчивая наш краткий обзор современного состояния и путей развития алгебры, мы должны еще раз подчеркнуть, что рассмотренные здесь вопросы в основном лежат за пределами курса высшей алгебры. Задачей обзора было лишь помочь читателю получить правильное представление о месте, занимаемом курсом высшей алгебры в алгебраической науке в целом и во всем большом здании математики.

ГЛАВА ПЕРВАЯ

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ. ОПРЕДЕЛИТЕЛИ

§ 1. Метод последовательного исключения неизвестных

Мы начинаем курс высшей алгебры с изучения систем уравнений первой степени с несколькими неизвестными или, как обычно говорят, *систем линейных уравнений* [1]).

Теория систем линейных уравнений кладет начало большому и важному отделу алгебры — линейной алгебре, — к которому относится большая часть глав нашей книги, в частности ее первые три главы. Коэффициенты уравнений, рассматриваемых в этих трех главах, значения неизвестных и вообще все числа, с которыми мы будем встречаться, следует считать действительными. Впрочем, все содержание этих глав дословно переносится и на случай произвольных комплексных чисел, уже известных читателю из курса средней школы.

В отличие от элементарной алгебры мы будем изучать системы с произвольным числом уравнений и неизвестных, причем иногда число уравнений системы не будет даже предполагаться совпадающим с числом неизвестных. Пусть нам дана система из s линейных уравнений с n неизвестными. Условимся употреблять следующую символику: неизвестные мы будем обозначать буквой x с индексами: $x_1, x_2, \ldots, x_n$; уравнения будем считать перенумерованными — первое, второе, ..., s-е; коэффициент из i-го уравнения при неизвестном x_j обозначим через a_{ij} [2]); наконец, свободный член i-го уравнения обозначим через b_i.

---

[1]) Это название связано с тем, что в аналитической геометрии уравнение первой степени с двумя неизвестными определяет прямую линию на плоскости.

[2]) Мы употребляем, следовательно, два индекса, из которых первый указывает на номер уравнения, второй — на номер неизвестного. Для сокращения письма эти индексы не разделяются запятой; не следует, однако, в случае a_{11} вместо «а один один» читать «а одиннадцать», в случае a_{34} вместо «а три четыре» читать «а тридцать четыре».

Наша система запишется теперь в следующем общем виде:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = b_s. \end{array}\right\} \quad (1)$$

Коэффициенты при неизвестных составляют прямоугольную таблицу

$$\begin{pmatrix} a_{11}a_{12} & \ldots & a_{1n} \\ a_{21}a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots \\ a_{s1}a_{s2} & \ldots & a_{sn} \end{pmatrix}, \quad (2)$$

называемую *матрицей* из *s строк* и *n столбцов*; числа a_{ij} называются *элементами* матрицы[1]). Если $s = n$ (т. е. число строк равно числу столбцов), то матрица называется *квадратной матрицей порядка n*. Диагональ этой матрицы, идущая от левого верхнего к первому нижнему углу (т. е. составленная из элементов $a_{11}, a_{22}, \ldots, a_{nn}$), называется *главной диагональю*. Квадратная матрица порядка n будет называться *единичной матрицей порядка n*, если все элементы ее главной диагонали равны единице, а все элементы вне этой диагонали равны нулю.

Решением системы линейных уравнений (1) называется такая система n чисел $k_1, k_2, \ldots, k_n$, что каждое из уравнений (1) обращается в тождество после замены в нем неизвестных x_i соответствующими числами k_i, $i = 1, 2, \ldots, n$[2]).

Система линейных уравнений может не иметь ни одного решения и тогда она называется *несовместной*. Такова, например, система

$$\begin{array}{l} x_1 + 5x_2 = 1, \\ x_1 + 5x_2 = 7; \end{array}$$

левые части этих уравнений совпадают, но правые различны, и поэтому никакая система значений неизвестных не может удовлетворить обоим уравнениям сразу.

Если же система линейных уравнений обладает решениями, то она называется *совместной*. Совместная система называется *определенной*, если она обладает одним-единственным решением — лишь такие системы допускаются к рассмотрению в элементарной алгебре,—

[1]) Таким образом, если матрицу (2) рассматривать вне связи с системой (1), то первый индекс элемента a_{ij} указывает на номер строки, второй — на номер столбца, на пересечении которых этот элемент стоит.

[2]) Подчеркиваем, что числа $k_1, k_2, \ldots, k_n$ составляют одно решение системы, а не n решений.

и ***неопределенной***, если решений больше чем одно; как мы узнаем позже, их будет в этом случае даже бесконечно много. Так, система

$$\left.\begin{aligned} x_1+2x_2&=7,\\ x_1+x_2&=4 \end{aligned}\right\}$$

определенна: она имеет решение $x_1=1$, $x_2=3$ и, как легко проверяется методом исключения неизвестного, это решение будет единственным. С другой стороны, система

$$\left.\begin{aligned} 3x_1-x_2&=1,\\ 6x_1-2x_2&=2 \end{aligned}\right\}$$

неопределенна, так как имеет бесконечно много решений вида

$$x_1=k, \qquad x_2=3k-1, \tag{3}$$

где число k произвольно, причем решениями, получающимися по формулам (3), исчерпываются все решения нашей системы.

Задача теории систем линейных уравнений состоит в разработке методов, позволяющих узнать, совместна ли данная система уравнений или нет, в случае совместности установить число решений, а также указать способ найти все эти решения.

Мы начнем с метода, наиболее удобного для практического разыскания решений систем с числовыми коэффициентами, а именно *метода последовательного исключения неизвестных* или *метода Гаусса*.

Вначале одно предварительное замечание. Нам придется делать дальше такие преобразования системы линейных уравнений: обе части одного из уравнений системы, умноженные на одно и то же число, вычитать из соответствующих частей некоторого другого уравнения системы. Пусть для определенности мы обе части первого уравнения системы (1), умноженные на число c, вычитаем из соответствующих частей второго уравнения. Мы получим новую систему линейных уравнений:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ \dots\ +a_{1n}x_n&=b_1,\\ a'_{21}x_1+a'_{22}x_2+\ \dots\ +a'_{2n}x_n&=b'_2,\\ a_{31}x_1+a_{32}x_2+\ \dots\ +a_{3n}x_n&=b_3,\\ \dots\dots\dots\dots\dots\dots\dots&\dots\\ a_{s1}x_1+a_{s2}x_2+\ \dots\ +a_{sn}x_n&=b_s, \end{aligned}\right\} \tag{4}$$

где

$$a'_{2j}=a_{2j}-ca_{1j} \quad \text{при} \quad j=1,\ 2,\ \dots,\ n, \quad b'_2=b_2-cb_1.$$

Системы уравнений (1) *и* (4) *эквивалентны, т. е. они или обе несовместны, или же обе совместны и обладают одними и теми же решениями.* В самом деле, пусть $k_1, k_2, \dots, k_n$ будет

произвольное решение системы (1). Эти числа удовлетворяют, очевидно, всем уравнениям системы (4), кроме второго. Они удовлетворяют, однако, и второму уравнению системы (4) — достаточно вспомнить, как это уравнение выражается через второе и первое уравнения системы (1). Обратно, всякое решение системы (4) будет удовлетворять и системе (1). Действительно, второе уравнение системы (1) получается вычитанием из обеих частей второго уравнения системы (4) соответствующих частей первого уравнения этой системы, умноженных на число $-c$.

Понятно, что *если к системе* (1) *будут несколько раз применены преобразования рассмотренного типа, то вновь полученная система уравнений останется эквивалентной исходной системе* (1).

Может случиться, что после выполнения таких преобразований в нашей системе появится уравнение, все коэффициенты левой части которого равны нулю. Если и свободный член этого уравнения равен нулю, то уравнение удовлетворяется при любых значениях неизвестных, и поэтому, *отбрасывая это уравнение, мы придем к системе уравнений, эквивалентной исходной системе.* Если же свободный член рассматриваемого уравнения отличен от нуля, то это уравнение не может удовлетворяться ни при каких значениях неизвестных, а поэтому *полученная нами система уравнений, равно как и эквивалентная ей исходная система, будут несовместными.*

Перейдем теперь к изложению метода Гаусса.

Пусть дана произвольная система линейных уравнений (1). Положим, для определенности, что коэффициент $a_{11} \neq 0$, хотя на самом деле он может, конечно, оказаться равным нулю, и мы должны будем начать с какого-либо другого, отличного от нуля, коэффициента из первого уравнения системы.

Преобразуем теперь систему (1), исключая неизвестное x_1 из всех уравнений, кроме первого. Для этого обе части первого уравнения умножим на число $\frac{a_{21}}{a_{11}}$ и вычтем из соответствующих частей второго уравнения, затем обе части первого уравнения, умноженные на число $\frac{a_{31}}{a_{11}}$, вычтем из соответствующих частей третьего уравнения, и т. д.

Мы придем этим путем к новой системе из s линейных уравнений с n неизвестными:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2, \\ a'_{32}x_2 + a'_{33}x_3 + \ldots + a'_{3n}x_n &= b'_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{s2}x_2 + a'_{s3}x_3 + \ldots + a'_{sn}x_n &= b'_s. \end{aligned}\right\} \tag{5}$$

Нам нет необходимости явно записывать выражения новых коэффициентов a'_{ij} и новых свободных членов b'_i через коэффициенты и свободные члены исходной системы (1).

Как мы знаем, система уравнений (5) эквивалентна системе (1). Будем преобразовывать теперь систему (5). При этом первое уравнение мы не будем больше трогать совсем и подлежащей преобразованиям будем считать лишь часть системы (5), состоящую из всех уравнений, кроме первого. При этом мы считаем, конечно, что среди этих уравнений нет таких, все коэффициенты левых частей которых равны нулю,— такие уравнения мы выбросили бы, если бы и их свободные члены были равны нулю, а в противном случае мы уже доказали бы несовместность нашей системы. Таким образом, среди коэффициентов a'_{ij} есть отличные от нуля; для определенности примем, что $a'_{22} \neq 0$. Преобразуем теперь систему (5), вычитая из обеих частей третьего и каждого из следующих уравнений обе части второго уравнения, умноженные соответственно на числа

$$\frac{a'_{32}}{a'_{22}},\ \frac{a'_{42}}{a'_{22}},\ \ldots,\ \frac{a'_{s2}}{a'_{22}}.$$

Этим будет исключено неизвестное x_2 из всех уравнений, кроме первого и второго, и мы придем к следующей системе уравнений, эквивалентной системе (5), а поэтому и системе (1):

$$\left.\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1,\\
a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2,\\
a''_{33}x_3 + \ldots + a''_{3n}x_n &= b''_3,\\
\ldots\ldots\ldots\ldots\ldots&\\
a''_{t3}x_3 + \ldots + a''_{tn}x_n &= b''_t.
\end{aligned}\right\}$$

Наша система содержит теперь t уравнений, $t \leqslant s$, так как некоторые уравнения оказались, возможно, отброшенными. Понятно, что число уравнений системы могло уменьшиться уже после исключения неизвестного x_1. В дальнейшем подлежит преобразованиям лишь часть полученной системы, содержащая все уравнения, кроме двух первых.

Когда остановится этот процесс последовательного исключения неизвестных?

Если мы придем к такой системе, одно из уравнений которой имеет отличный от нуля свободный член, а все коэффициенты левой части равны нулю, то, как мы знаем, наша исходная система несовместна.

В противном случае мы получим следующую систему уравнений, эквивалентную системе (1):

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,\,k-1}x_{k-1} + a_{1k}x_k + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + \ldots + a'_{2,k-1}x_{k-1} + a'_{2k}x_k + \ldots + a'_{2n}x_n &= b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a^{(k-2)}_{k-1,\,k-1}x_{k-1} + a^{(k-2)}_{k-1,\,k}x_k + \ldots + a^{(k-2)}_{k-1,\,n}x_n &= b^{(k-2)}_{k-1}, \\ a^{(k-1)}_{kk}x_k + \ldots + a^{(k-1)}_{kn}x_n &= b^{(k-1)}_k. \end{aligned}\right\} \quad (6)$$

Здесь $a_{11} \neq 0$, $a'_{22} \neq 0$, ..., $a^{(k-2)}_{k-1,\,k-1} \neq 0$, $a^{(k-1)}_{kk} \neq 0$. Отметим также, что $k \leqslant s$ и, очевидно, $k \leqslant n$.

В этом случае система (1) *совместная. Она будет определенной при $k=n$ и неопределенной при $k<n$.*

Действительно, если $k=n$, то система (6) имеет вид

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a^{(n-1)}_{nn}x_n &= b^{(n-1)}_n. \end{aligned}\right\} \quad (7)$$

Из последнего уравнения мы получаем вполне определенное значение для неизвестного x_n. Подставляя его в предпоследнее уравнение, мы найдем однозначно определенное значение для неизвестного x_{n-1}. Продолжая так далее, мы найдем, что система (7), а поэтому и система (1), обладают единственным решением, т. е. совместны и определенны.

Если же $k<n$, то для «свободных» неизвестных $x_{k+1}, \ldots, x_n$ мы возьмем произвольные числовые значения, после чего, двигаясь по системе (6) снизу вверх, мы, как и выше, найдем для неизвестных $x_k, x_{k-1}, \ldots, x_2, x_1$ вполне определенные значения. Так как значения для свободных неизвестных можно выбрать бесконечным числом различных способов, то наша система (6) и, следовательно, система (1) будут совместными, но неопределенными. Легко проверить, что указанным здесь методом (при всевозможных выборах значений для свободных неизвестных) будут найдены все решения системы (1).

В первый момент кажется возможным еще один вид, к которому может приводиться система линейных уравнений методом Гаусса, а именно вид, получающийся приписыванием к системе (7) еще нескольких уравнений, содержащих лишь неизвестное x_n. В действительности, однако, в этом случае преобразования просто не доведены до конца: так как $a^{(n-1)}_{nn} \neq 0$, то из всех уравнений, начиная с $(n+1)$-го, неизвестное x_n может быть исключено.

Следует заметить, что «треугольная» форма системы уравнений (7) или «трапецоидальная» форма системы уравнений (6) (при $k<n$)

получилась ввиду предположения, что коэффициенты a_{11}, a'_{22} и т. д. отличны от нуля. В общем случае та система уравнений, к которой мы придем после доведения до конца процесса исключения неизвестных, приобретет треугольную или трапецоидальную форму лишь после надлежащего изменения нумерации неизвестных.

Суммируя все изложенное выше, мы получаем, что *метод Гаусса применим к любой системе линейных уравнений. При этом система будет несовместной, если в процессе преобразований мы получим уравнение, в котором коэффициенты при всех неизвестных равны нулю, а свободный член отличен от нуля; если же мы такого уравнения не встретим, то система будет совместной. Совместная система уравнений будет определенной, если она приводится к треугольному виду* (7), *и неопределенной, если приводится к трапецоидальному виду* (6) *при* $k<n$.

Применим сказанное к случаю системы линейных *однородных* уравнений, т. е. уравнений, свободные члены которых равны нулю. Такая система всегда совместна, так как обладает *нулевым решением* (0, 0, ..., 0). Пусть в рассматриваемой системе число уравнений меньше числа неизвестных. Тогда наша система не может приводиться к треугольному виду, так как в процессе преобразований по методу Гаусса число уравнений системы может уменьшаться, но не может увеличиваться; она приводится, следовательно, к трапецоидальному виду, т. е. неопределенна.

Иными словами, *если в системе линейных однородных уравнений число уравнений меньше числа неизвестных, то эта система обладает, помимо нулевого решения, также и ненулевыми решениями*, т. е. решениями, в которых значения некоторых (или даже всех) неизвестных отличны от нуля; *таких решений будет бесконечно много.*

При практическом решении системы линейных уравнений методом Гаусса следует выписать матрицу из коэффициентов системы, присоединить к ней столбец из свободных членов, для удобства отделенный вертикальной чертой, и все преобразования выполнять над строками этой «расширенной» матрицы.

Примеры. 1. Решить систему.

$$\left.\begin{aligned} x_1+2x_2+5x_3&=-9,\\ x_1-x_2+3x_3&=2,\\ 3x_1-6x_2-x_3&=25. \end{aligned}\right\}$$

Подвергаем преобразованиям расширенную матрицу этой системы:

$$\left(\begin{array}{ccc|c} 1 & 2 & 5 & -9\\ 1 & -1 & 3 & 2\\ 3 & -6 & -1 & 25 \end{array}\right)\to\left(\begin{array}{ccc|c} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & -12 & -16 & 52 \end{array}\right)\to\left(\begin{array}{ccc|c} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & 0 & -8 & 8 \end{array}\right).$$

Мы приходим, следовательно, к системе уравнений

$$\left.\begin{aligned} x_1+2x_2+5x_3&=-9,\\ -3x_2-2x_3&=11,\\ -8x_3&=8,\end{aligned}\right\}$$

обладающей единственным решением

$$x_1=2,\quad x_2=-3,\quad x_3=-1.$$

Исходная система оказалась определенной.

2. Решить систему

$$\left.\begin{aligned} x_1-5x_2-8x_3+x_4&=3,\\ 3x_1+x_2-3x_3-5x_4&=1,\\ x_1\qquad-7x_3+2x_4&=-5,\\ 11x_2+20x_3-9x_4&=2.\end{aligned}\right\}$$

Преобразуем расширенную матрицу системы:

$$\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 3 & 1 & -3 & -5 & 1\\ 1 & 0 & -7 & 2 & -5\\ 0 & 11 & 20 & -9 & 2\end{array}\right)\rightarrow\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & 16 & 21 & -8 & -8\\ 0 & 5 & 1 & 1 & -8\\ 0 & 11 & 20 & -9 & 2\end{array}\right)\rightarrow$$

$$\rightarrow\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & -89 & 0 & -29 & 162\end{array}\right)\rightarrow\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & 0 & 0 & 0 & 2\end{array}\right).$$

Мы пришли к системе, содержащей уравнение $0=2$. Исходная система будет, следовательно, несовместной.

3. Решить систему

$$\left.\begin{aligned} 4x_1+x_2-3x_3-x_4&=0,\\ 2x_1+3x_2+x_3-5x_4&=0,\\ x_1-2x_2-2x_3+3x_4&=0.\end{aligned}\right\}$$

Это система однородных уравнений, причем число уравнений меньше числа неизвестных; она должна быть поэтому неопределенной. Так как все свободные члены равны нулю, то мы будем подвергать преобразованиям лишь матрицу из коэффициентов системы:

$$\begin{pmatrix} 4 & 1 & -3 & -1\\ 2 & 3 & 1 & -5\\ 1 & -2 & -2 & 3\end{pmatrix}\rightarrow\begin{pmatrix} 0 & 9 & 5 & -13\\ 0 & 7 & 5 & -11\\ 1 & -2 & -2 & 3\end{pmatrix}\rightarrow\begin{pmatrix} 0 & 2 & 0 & -2\\ 0 & 7 & 5 & -11\\ 1 & -2 & -2 & 3\end{pmatrix}.$$

Мы пришли к системе уравнений

$$\left.\begin{aligned} 2x_2\qquad-2x_4&=0,\\ 7x_2+5x_3-11x_4&=0,\\ x_1-2x_2-2x_3+3x_4&=0.\end{aligned}\right\}$$

В качестве свободного неизвестного можно принять любое из неизвестных x_2 и x_4. Пусть $x_4=\alpha$. Тогда из первого уравнения следует $x_2=\alpha$, после чего

из второго уравнения получаем $x_3=\frac{4}{5}\alpha$ и, наконец, из третьего уравнения $x_1=\frac{3}{5}\alpha$. Таким образом,

$$\frac{3}{5}\alpha,\ \alpha,\ \frac{4}{5}\alpha,\ \alpha$$

будет общий вид решений заданной системы уравнений.

§ 2. Определители второго и третьего порядков

Метод решения системы линейных уравнений, изложенный в предшествующем параграфе, весьма прост и требует выполнения однотипных вычислений, легко осуществляемых на счетных машинах. Его существенным недостатком является, однако, то, что он не дает возможности сформулировать условия совместности или определенности системы при помощи коэффициентов и свободных членов этой системы. С другой стороны, даже в случае определенной системы этот метод не позволяет найти формулы, выражающие решение системы через ее коэффициенты и свободные члены. Все это оказывается, однако, необходимым в разных теоретических вопросах, в частности, в геометрических исследованиях, а поэтому теорию систем линейных уравнений приходится развивать иными методами, более глубокими. Общий случай будет рассмотрен в следующей главе, а дальнейшее содержание настоящей главы посвящается случаю определенных систем, имеющих равное число уравнений и неизвестных, причем мы начнем с уже изучавшихся в элементарной алгебре систем с двумя и тремя неизвестными.

Пусть дана *система двух линейных уравнений с двумя неизвестными*

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2&=b_1,\\ a_{21}x_1+a_{22}x_2&=b_2,\end{aligned}\right\} \tag{1}$$

коэффициенты которой составляют квадратную матрицу второго порядка

$$\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}. \tag{2}$$

Применяя к системе (1) метод уравнивания коэффициентов, мы получим:

$$(a_{11}a_{22}-a_{12}a_{21})\,x_1=b_1a_{22}-a_{12}b_2,$$
$$(a_{11}a_{22}-a_{12}a_{21})\,x_2=a_{11}b_2-b_1a_{21}.$$

Предположим, что $a_{11}a_{22}-a_{12}a_{21}\neq 0$. Тогда

$$x_1=\frac{b_1a_{22}-a_{12}b_2}{a_{11}a_{22}-a_{12}a_{21}},\quad x_2=\frac{a_{11}b_2-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}. \tag{3}$$

Легко проверить, подставляя полученные значения неизвестных в уравнения (1), что (3) служит решением для системы (1); вопрос о единственности этого решения будет рассматриваться в § 7.

Общий знаменатель значений неизвестных (3) очень просто выражается через элементы матрицы (2): он равен произведению элементов главной диагонали минус произведение элементов второй диагонали. Это число называется *определителем* (или *детерминантом*) матрицы (2), причем, как говорят, *определителем второго порядка*, так как матрица (2) есть матрица второго порядка. Для обозначения определителя матрицы (2) употребляется следующий символ: выписывается матрица (2), но заключается вместо круглых скобок в прямые черточки; таким образом,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{4}$$

Примеры.

1) $$\begin{vmatrix} 3 & 7 \\ 1 & 4 \end{vmatrix} = 3\cdot 4 - 7\cdot 1 = 5;$$

2) $$\begin{vmatrix} 1 & -2 \\ 3 & 5 \end{vmatrix} = 1\cdot 5 - (-2)\cdot 3 = 11.$$

Следует еще раз подчеркнуть, что в то время как матрица есть таблица из чисел, определитель есть число, вполне определенным образом связанное с квадратной матрицей. Заметим, что произведения $a_{11}a_{22}$ и $a_{12}a_{21}$ называются *членами* определителя второго порядка.

Числители выражений (3) имеют такой же вид, как и знаменатель, т. е. также являются определителями второго порядка: числитель выражения для x_1 является определителем матрицы, получающейся из матрицы (2) заменой ее первого столбца столбцом из свободных членов системы (1), числитель выражения для x_2 есть определитель матрицы, получающейся из матрицы (2) такой же заменой ее второго столбца. Формулы (3) теперь можно записать в следующем виде:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{5}$$

Словами это правило решения системы двух линейных уравнений с двумя неизвестными (называемое *правилом Крамера*) формулируется следующим образом:

Если определитель (4) *из коэффициентов системы уравнений* (1) *отличен от нуля, то мы получим решение системы* (1), *беря в качестве значений для неизвестных дроби, общим знаменателем которых служит определитель* (4), *а числителем для неизвестного* x_i $(i = 1, 2)$ *является определитель, получающийся заменой*

в определителе (4) *i-го столбца* (*т. е. столбца коэффициентов при искомом неизвестном*) *столбцом из свободных членов системы* (1)[1].

Пример. Решить систему

$$\left.\begin{aligned} 2x_1 + x_2 &= 7, \\ x_1 - 3x_2 &= -2. \end{aligned}\right\}$$

Определитель из коэффициентов есть

$$d = \begin{vmatrix} 2 & 1 \\ 1 & -3 \end{vmatrix} = -7;$$

он отличен от нуля, и поэтому к системе применимо правило Крамера. Числителями для неизвестных будут определители

$$d_1 = \begin{vmatrix} 7 & 1 \\ -2 & -3 \end{vmatrix} = -19, \qquad d_2 = \begin{vmatrix} 2 & 7 \\ 1 & -2 \end{vmatrix} = -11.$$

Таким образом, решением нашей системы служит следующая система чисел:

$$x_1 = \frac{d_1}{d} = \frac{19}{7}, \qquad x_2 = \frac{d_2}{d} = \frac{11}{7}.$$

Введение определителей второго порядка не вносит существенных упрощений в решение систем двух линейных уравнений с двумя неизвестными, и без этого не представляющее никаких трудностей. Аналогичные методы для случая *систем трех линейных уравнений с тремя неизвестными* оказываются уже практически полезными. Пусть дана система

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}\right\} \tag{6}$$

с матрицей из коэффициентов

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{7}$$

Если мы умножим обе части первого из уравнений (6) на число $a_{22}a_{33} - a_{23}a_{32}$, обе части второго уравнения на $a_{13}a_{32} - a_{12}a_{33}$, обе части третьего на $a_{12}a_{23} - a_{13}a_{22}$, а затем сложим все три уравнения, то, как легко проверить, коэффициенты при x_2 и x_3 окажутся

[1]) Мы в этой формулировке для краткости говорим о замене столбцов «в определителе». Подобным же образом и в будущем мы будем говорить, если это будет удобнее, о строках и столбцах определителя, о его элементах, диагоналях и т. д.

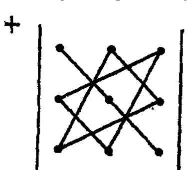
равными нулю, т. е. эти неизвестные одновременно исключаются, и мы получим равенство

$$(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-\\ -a_{11}a_{23}a_{32})\,x_1=b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-a_{13}a_{22}b_3-\\ -a_{12}b_2a_{33}-b_1a_{23}a_{32}. \qquad (8)$$

Коэффициент при x_1 в этом равенстве называется *определителем третьего порядка*, соответствующим матрице (7). Для его записи употребляется такая же символика, как и в случае определителей второго порядка; таким образом,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-\\ -a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}. \qquad (9)$$

Хотя выражение определителя третьего порядка является достаточно громоздким, закон его составления из элементов матрицы (7) оказывается весьма простым. В самом деле, один из трех членов определителя, входящих в его выражение (9) со знаком плюс, будет произведением элементов главной диагонали, каждый из двух других — произведением элементов, лежащих на параллели к этой диагонали, с добавлением третьего множителя из противоположного угла матрицы. Члены, входящие в (9) со знаком минус, строятся таким же образом, но относительно второй диагонали. Мы получаем способ вычисления определителей третьего порядка, приводящий (при наличии некоторой тренировки) весьма быстро к результату. На рис. 1



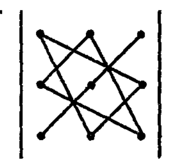


Рис. 1.

слева схематически указано правило вычисления положительных членов определителя третьего порядка, справа — правило вычисления его отрицательных членов.

Примеры.

1)
$$\begin{vmatrix} 2 & 1 & 2 \\ -4 & 3 & 1 \\ 2 & 3 & 5 \end{vmatrix} = 2\cdot3\cdot5+1\cdot1\cdot2+2\cdot(-4)\cdot3-\\ -2\cdot3\cdot2-1\cdot(-4)\cdot5-2\cdot1\cdot3=\\ =30+2-24-12+20-6=10.$$

2)
$$\begin{vmatrix} 1 & 0 & -5 \\ -2 & 3 & 2 \\ 1 & -2 & 0 \end{vmatrix} = 1\cdot3\cdot0+0\cdot2\cdot1+(-5)\cdot(-2)\cdot(-2)-\\ -(-5)\cdot3\cdot1-0\cdot(-2)\cdot0-1\cdot2\cdot(-2)=\\ =-20+15+4=-1.$$

Правая часть равенства (8) также будет определителем третьего порядка, а именно определителем матрицы, получающейся из матрицы (7) заменой ее первого столбца столбцом из свободных членов системы (6). Если мы обозначим определитель (9) буквой d, а определитель, получающийся заменой его j-го столбца ($j=1, 2, 3$) столбцом из свободных членов системы (6), символом d_j, то равенство (8) приобретет вид $dx_1=d_1$, откуда при $d\neq 0$ следует

$$x_1=\frac{d_1}{d}. \qquad (10)$$

Таким же путем, умножая уравнения (6) соответственно на числа $a_{23}a_{31}-a_{21}a_{33}$, $a_{11}a_{33}-a_{13}a_{31}$, $a_{13}a_{21}-a_{11}a_{23}$, мы получим для x_2 следующее выражение (снова при $d\neq 0$):

$$x_2=\frac{d_2}{d}. \qquad (11)$$

Наконец, умножая эти уравнения соответственно на $a_{21}a_{32}-a_{22}a_{31}$, $a_{12}a_{31}-a_{11}a_{32}$, $a_{11}a_{22}-a_{12}a_{21}$, мы придем к выражению для x_3:

$$x_3=\frac{d_3}{d}. \qquad (12)$$

Подставляя выражения (10)—(12) в уравнения (6) (предполагается, понятно, что определители d и все d_j записаны в развернутом виде), мы получили бы после громоздких, но вполне доступных читателю вычислений, что все эти уравнения удовлетворяются, т. е. что числа (10)—(12) составляют решение системы (6). Таким образом, *если определитель из коэффициентов системы трех линейных уравнений с тремя неизвестными отличен от нуля, то решение этой системы может быть найдено по правилу Крамера, формулируемому так же, как и в случае системы двух уравнений.* Другое доказательство этого утверждения (не опирающееся на опущенные нами вычисления), а также доказательство единственности решения (10)—(12) системы (6), притом для более общего случая, читатель найдет в § 7.

Пример. Решить систему:

$$\left.\begin{aligned} 2x_1-\ x_2+\ x_3&=0,\\ 3x_1+2x_2-5x_3&=1,\\ x_1+3x_2-2x_3&=4. \end{aligned}\right\}$$

Определитель из коэффициентов системы отличен от нуля:

$$d=\begin{vmatrix} 2 & -1 & 1\\ 3 & 2 & -5\\ 1 & 3 & -2 \end{vmatrix}=28,$$

поэтому к системе применимо правило Крамера. Числителями для неизвестных будут определители

$$d_1=\begin{vmatrix} 0 & -1 & 1 \\ 1 & 2 & -5 \\ 4 & 3 & -2 \end{vmatrix}=13, \quad d_2=\begin{vmatrix} 2 & 0 & 1 \\ 3 & 1 & -5 \\ 1 & 4 & -2 \end{vmatrix}=47,$$

$$d_3=\begin{vmatrix} 2 & -1 & 0 \\ 3 & 2 & 1 \\ 1 & 3 & 4 \end{vmatrix}=21,$$

т. е. решением системы служит система чисел

$$x_1=\frac{13}{28}, \quad x_2=\frac{47}{28}, \quad x_3=\frac{21}{28}=\frac{3}{4}.$$

§ 3. Перестановки и подстановки

Для определения и изучения определителей порядка n нам будут нужны некоторые понятия и факты, относящиеся к конечным множествам. Пусть дано некоторое конечное множество M, состоящее из n элементов. Эти элементы могут быть перенумерованы при помощи первых n натуральных чисел $1, 2, \ldots, n$, и так как в интересующих нас вопросах индивидуальные свойства элементов множества M не будут играть никакой роли, то мы просто примем, что элементами M служат сами эти числа $1, 2, \ldots, n$.

Помимо употребляющегося нами расположения чисел $1, 2, \ldots, n$ в их нормальном порядке, их можно упорядочить и многими другими способами. Так, числа 1, 2, 3, 4 можно расположить также следующими способами: 3, 1, 2, 4 или 2, 4, 1, 3 и т. д. Всякое расположение чисел $1, 2, \ldots, n$ в некотором определенном порядке называется *перестановкой* из n чисел (или из n символов).

Число различных перестановок из n символов равно произведению $1\cdot2\ldots n$, обозначаемому $n!$ (читается: «эн-факториал»). Действительно, общий вид перестановки из n символов есть $i_1, i_2, \ldots, i_n$, где каждое из i_s есть одно из чисел $1, 2, \ldots, n$, причем ни одно из этих чисел не встречается дважды. В качестве i_1 можно взять любое из чисел $1, 2, \ldots, n$; это дает n различных возможностей. Если, однако, i_1 уже выбрано, то в качестве i_2 можно взять лишь одно из оставшихся $n-1$ чисел, т. е. число различных способов выбрать символы i_1 и i_2 равно произведению $n(n-1)$ и т. д.

Таким образом, число перестановок из n символов при $n=2$ равно $2!=2$ (перестановки 12 и 21; мы не будем в примерах, где $n\leqslant 9$, разделять переставляемые символы запятыми); при $n=3$ это число равно $3!=6$, при $n=4$ оно равно $4!=24$. Далее, с ростом n число перестановок чрезвычайно быстро возрастает; так, при $n=5$ оно равно $5!=120$, а при $n=10$ — уже 3 628 800.

Если в некоторой перестановке мы поменяем местами какие-либо два символа (не обязательно стоящие рядом), а все остальные сим-

волы оставим на месте, то получим, очевидно, новую перестановку. Это преобразование перестановки называется *транспозицией*.

Все $n!$ перестановок из n символов можно расположить в таком порядке, что каждая следующая будет получаться из предыдущей одной транспозицией, причем начинать можно с любой перестановки.

Это утверждение справедливо при $n=2$: если требуется начинать с перестановки 12, то искомое расположение будет 12, 21; если же мы должны начать с перестановки 21, то это будет расположение 21, 12. Предположим, что наше утверждение уже доказано для $n-1$, и докажем его для n. Пусть мы должны начать с перестановки

$$i_1, i_2, \ldots, i_n. \tag{1}$$

Рассмотрим все перестановки из n символов, у которых на первом месте стоит i_1. Таких перестановок $(n-1)!$ и их можно упорядочить в согласии с требованиями теоремы, притом начиная с перестановки (1), так как это сводится на самом деле к упорядочению всех перестановок из $n-1$ символов, которое, по индуктивному предположению, можно начать с любой перестановки, в частности с перестановки $i_2, \ldots, i_n$. В последней из полученных таким путем перестановок из n символов совершаем транспозицию символа i_1 с любым другим символом, например с i_2, и, начиная с вновь полученной перестановки, упорядочиваем нужным образом все те перестановки, у которых на первом месте стоит i_2, и т. д. Этим путем можно, очевидно, перебрать все перестановки из n символов.

Из этой теоремы вытекает, что *от любой перестановки из n символов можно перейти к любой другой перестановке из тех же символов при помощи нескольких транспозиций.*

Говорят, что в данной перестановке числа i и j составляют *инверсию*, если $i>j$, но i стоит в этой перестановке раньше j. Перестановка называется *четной*, если ее символы составляют четное число инверсий, и *нечетной*—в противоположном случае. Так, перестановка $1, 2, \ldots, n$ будет четной при любом n, так как число инверсий в ней равно нулю. Перестановка 451362 ($n=6$) содержит 8 инверсий и поэтому четная. Перестановка 38524671 ($n=8$) содержит 15 инверсий и поэтому нечетная.

Всякая транспозиция меняет четность перестановки.

Для доказательства этой важной теоремы рассмотрим сначала случай, когда транспонируемые символы i и j стоят рядом, т. е. перестановка имеет вид $\ldots, i, j, \ldots$, где многоточия заменяют те символы, которые не затрагиваются транспозицией. Транспозиция превращает нашу перестановку в перестановку $\ldots, j, i, \ldots$, причем, понятно, в обеих перестановках каждый из символов i, j составляет одни и те же инверсии с символами, остающимися на месте. Если символы i и j раньше не составляли инверсии, то

в новой перестановке появляется одна новая инверсия, т. е. число инверсий увеличивается на единицу; если же они раньше составляли инверсию, то теперь она пропадает, т. е. число инверсий на единицу уменьшается. В обоих случаях четность перестановки меняется.

Пусть теперь между транспонируемыми символами i и j расположены s символов, $s>0$, т. е. перестановка имеет вид

$$\ldots, i, k_1, k_2, \ldots, k_s, j, \ldots \qquad (2)$$

Транспозицию символов i и j можно получить в результате последовательного выполнения $2s+1$ транспозиций соседних элементов. А именно, это будут транспозиции, переставляющие символы i и k_1, затем i (уже стоящие на месте символа k_1) и k_2 и т. д., пока i не займет место символа k_s. За этими s транспозициями следует транспозиция, перемещающая символы i и j, а затем s транспозиций символа j со всеми k, после чего j занимает место символа i, а символы k возвращаются на свои старые места. Таким образом, мы нечетное число раз меняли четность перестановки, а поэтому перестановки (2) и

$$\ldots, j, k_1, k_2, \ldots, k_s, i, \ldots \qquad (3)$$

имеют противоположные четности.

При $n \geqslant 2$ *число четных перестановок из n символов равно числу нечетных, т. е. равно* $\frac{1}{2}n!$.

В самом деле, упорядочим, на основании доказанного ранее, все перестановки из n символов так, что каждая получается из предыдущей одной транспозицией. Соседние перестановки будут иметь противоположные четности, т. е. перестановки расположены так, что четные и нечетные чередуются. Наше утверждение вытекает теперь из очевидного замечания, что при $n \geqslant 2$ число $n!$ четно.

Определим теперь одно новое понятие, а именно понятие *подстановки n-й степени*. Запишем одну под другой две перестановки из n символов, беря полученные две строки в скобки; например, при $n=5$:

$$\begin{pmatrix} 3 & 5 & 1 & 4 & 2 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix}. \qquad (4)$$

В этом примере[1]) под числом 3 стоит число 5, под числом 5 — число 2 и т. д. Мы скажем, что число 3 *переходит* в 5, число 5 переходит в 2, число 1 переходит в 3, число 4 переходит в 4 (или *остается на месте*) и, наконец, число 2 переходит в 1. Таким образом, две перестановки, записанные друг под другом в виде (4), определяют некоторое *взаимно однозначное отображение* множества из первых пяти натуральных чисел на себя, т. е. отображе-

---

[1]) Внешне он напоминает матрицу из двух строк и 5 столбцов, но имеет совсем иной смысл.

ние, которое каждому из натуральных чисел 1, 2, 3, 4, 5 ставит в соответствие одно из этих же натуральных чисел, причем разным числам ставятся в соответствие различные же числа. При этом, так как чисел всего пять, т. е. конечное множество, каждое из этих пяти чисел будет соответствовать одному из чисел 1, 2, 3, 4, 5, а именно числу, которое в него «переходит».

Ясно, что то взаимно однозначное отображение множества из первых пяти натуральных чисел, которое мы получили при помощи (4), можно было бы получить, записывая одну под другой и некоторые другие пары перестановок из пяти символов. Эти записи получаются из (4) путем нескольких транспозиций столбиков; таковы, например,

$$\begin{pmatrix} 2 & 1 & 5 & 3 & 4 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 2 & 4 & 3 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 5 & 1 & 4 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}. \tag{5}$$

Во всех этих записях 3 переходит в 5, 5 в 2, и т. д.

Аналогичным путем две перестановки из n символов, записанные одна под другой, определяют некоторое взаимно однозначное отображение множества первых n натуральных чисел на себя. Всякое взаимно однозначное отображение A множества первых n натуральных чисел на себя называется ***подстановкой n-й степени***, причем, очевидно, всякая подстановка A может быть записана при помощи двух перестановок, подписанных одна под другой

$$A = \begin{pmatrix} i_1, & i_2, & \ldots, & i_n \\ \alpha_{i_1}, & \alpha_{i_2}, & \ldots, & \alpha_{i_n} \end{pmatrix}; \tag{6}$$

через α_i здесь обозначается то число, в которое при подстановке A переходит число i, $i = 1, 2, \ldots, n$.

Подстановка A обладает многими различными записями вида (6). Так, (4) и (5) являются различными записями одной и той же подстановки 5-й степени.

От одной записи подстановки A к другой можно перейти при помощи нескольких транспозиций столбиков. При этом можно получить такую запись вида (6), в верхней (или нижней) строке которой стоит любая наперед заданная перестановка из n символов. В частности, всякая подстановка n-й степени A может быть записана в виде

$$A = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}, \tag{7}$$

т. е. с натуральным расположением чисел в верхней строке. При такой записи различные подстановки отличаются друг от друга перестановками, стоящими в нижней строке, и поэтому *число подстановок n-й степени равно числу перестановок из n символов, т. е. равно n!*

Примером подстановки n-й степени служит *тождественная подстановка*

$$E=\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix},$$

при которой на месте остаются все символы.

Следует заметить, что верхняя и нижняя строки в записи (6) подстановки A играют разные роли и, переставив их, мы, вообще говоря, получаем другую подстановку. Так, подстановки 4-й степени

$$\begin{pmatrix}2 & 1 & 4 & 3\\ 4 & 3 & 1 & 2\end{pmatrix} \text{ и } \begin{pmatrix}4 & 3 & 1 & 2\\ 2 & 1 & 4 & 3\end{pmatrix}$$

различны: при первой число 2 переходит в 4, при второй — в 3.

Возьмем произвольную запись (6) некоторой подстановки n-й степени A. Перестановки, составляющие верхнюю и нижнюю строки этой записи, могут иметь или одинаковые, или противоположные четности. Переход к любой другой записи подстановки A можно осуществить, как мы знаем, путем последовательного выполнения нескольких транспозиций в верхней строке и соответствующих им транспозиций в нижней строке. Однако, совершая одну транспозицию в верхней строке записи (6) и одну транспозицию соответствующих элементов в нижней строке, мы одновременно меняем четности обеих перестановок и поэтому сохраняем совпадение или противоположность этих четностей. Отсюда следует, что *либо при всех записях подстановки A четности верхней и нижней строк совпадают, либо же при всех записях они противоположны.* В первом случае подстановка A будет называться *четной*, во втором — *нечетной*. В частности, тождественная подстановка будет четной.

Если подстановка A записана в виде (7), т. е. в верхней строке стоит четная перестановка $1, 2, \dots, n$, то четность подстановки A будет определяться четностью перестановки $\alpha_1, \alpha_2, \dots, \alpha_n$, стоящей в нижней строке. Отсюда следует, что *число четных подстановок n-й степени равно числу нечетных, т. е. равно* $\frac{1}{2}n!$.

Определению четности подстановок можно дать следующую несколько измененную форму. Если в записи (6) четности обеих строк совпадают, то число инверсий или в обеих строках четное, или в обеих нечетное, т. е. общее число инверсий в двух строках записи (6) будет четным; если же четности строк записи (6) противоположны, то общее число инверсий в этих двух строках нечетно. Таким образом, *подстановка A будет четной, если общее число инверсий в двух строках любой ее записи четно, и нечетной — в противоположном случае.*

Пример. Пусть дана подстановка пятой степени

$$\begin{pmatrix} 3 & 1 & 4 & 5 & 2 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}.$$

В ее верхней строке 4 инверсии, в нижней 7 иверсий. Общее число инверсий в двух строках есть 11, и поэтому подстановка нечетна.

Перепишем эту подстановку в виде

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 4 & 3 \end{pmatrix}.$$

Число инверсий в верхней строке есть 0, в нижней 5, т. е. общее число снова нечетио. Мы видим, что при разных записях подстановки сохраняется четность общего числа инверсий, но не само это число.

Мы хотим указать теперь другие формы определения четности подстановок, эквивалентные приведенным выше[1]). Для этой цели определим *умножение подстановок*, представляющее и само по себе очень большой интерес. Подстановка n-й степени есть, как мы знаем, взаимно однозначное отображение множества чисел 1, 2, ..., n на себя. Результат последовательного выполнения двух взаимно однозначных отображений множества 1, 2, ..., n на себя снова будет, очевидно, некоторым взаимно однозначным отображением этого множества на себя, т. е. последовательное выполнение двух подстановок n-й степени приводит к некоторой вполне определенной третьей подстановке n-й степени, называемой *произведением* первой из заданных подстановок на вторую. Так, если даны подстановки четвертой степени

$$A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix},$$

то

$$AB=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

Действительно, при подстановке A символ 1 переходит в 3, но при B символ 3 переходит в 4, поэтому при AB символ 1 переходит в 4, и т. д.

Можно перемножить лишь подстановки одинаковой степени. *Умножение подстановок n-й степени при $n \geqslant 3$ некоммутативно.* Действительно, для рассмотренных выше подстановок A и B произведение BA имеет вид

$$BA=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix},$$

т. е. подстановка BA отлична от подстановки AB. Такие примеры можно подобрать для всех n при $n \geqslant 3$, хотя для некоторых пар подстановок закон коммутативности случайно может выполняться.

---

[1]) Они потребуются нам лишь в главе 14, и поэтому при первом чтении этот материал можно опустить.

Умножение подстановок ассоциативно, т. е. можно говорить о произведении любого конечного числа подстановок n-й степени, взятых (ввиду некоммутативности) в определенном порядке. В самом деле, пусть даны подстановки A, B и C и пусть символ i_1, $1 \leqslant i_1 \leqslant n$, переходит при подстановке A в символ i_2, i_2 при подстановке B переходит в символ i_3, а последний при подстановке C — в символ i_4. Тогда при подстановке AB символ i_1 переходит в i_3, при подстановке BC символ i_2 переходит в i_4, а поэтому как при $(AB)C$, так и при $A(BC)$ символ i_1 будет переходить в символ i_4.

Очевидно, что *произведение любой подстановки A на тождественную подстановку E, а также произведение E на A, равно A*:

$$AE = EA = A.$$

Назовем, наконец, *обратной* для подстановки A такую подстановку A^{-1} той же степени, что

$$AA^{-1} = A^{-1}A = E.$$

Легко видеть, что обратной для подстановки

$$A = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$$

служит подстановка

$$A^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ 1 & 2 & \ldots & n \end{pmatrix},$$

получающаяся из A переменой мест верхней и нижней строк.

Рассмотрим теперь подстановки специального вида, получающиеся из тождественной подстановки E при помощи одной транспозиции, производимой в ее нижней строке. Такие подстановки нечетны: они называются *транспозициями* и имеют вид

$$\begin{pmatrix} \ldots & i & \ldots & j & \ldots \\ \ldots & j & \ldots & i & \ldots \end{pmatrix}, \tag{8}$$

где многоточиями заменены символы, остающиеся на месте. Условимся обозначать эту транспозицию символом (i, j). Применение транспозиции символов i, j к нижней строке записи (7) произвольной подстановки A равносильно умножению подстановки A справа на подстановку (8), т. е. на (i, j). Мы знаем, что все перестановки из n символов можно получить из одной из них, например из $1, 2, \ldots, n$, последовательным выполнением транспозиций; поэтому всякая подстановка может быть получена из тождественной подстановки путем последовательного выполнения нескольких транспозиций в нижней строке, т. е. путем последовательного умножения на подстановки вида (8). Можно утверждать, следовательно (опуская множитель E), что *всякая подстановка представима в виде произведения транспозиций*.

Всякую подстановку можно многими разными способами разложить в произведение транспозиций. Всегда можно, например, добавить два одинаковых множителя вида (i, j) (i, j), которые дают в произведении подстановку E, т. е. взаимно уничтожаются. Укажем менее тривиальный пример:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} = (12)(15)(34) = (14)(24)(45)(34)(13).$$

Новый способ определения четности подстановки основан на следующей теореме:

При всех разложениях подстановки в произведение транспозиций четность числа этих транспозиций будет одна и та же, причем она совпадает с четностью самой подстановки.

Так, подстановка в рассмотренном выше примере будет нечетной, как можно проверить и подсчетом числа инверсий.

Эта теорема будет доказана, если мы покажем, что *произведение любых k транспозиций есть подстановка, четность которой совпадает с четностью числа k.* При $k=1$ это верно, так как транспозиция есть нечетная подстановка. Пусть наше утверждение уже доказано для случая $k-1$ множителей. Тогда его справедливость для k множителей вытекает из того, что числа $k-1$ и k имеют противоположные четности, а умножение подстановки (в данном случае — произведения первых $k-1$ множителей) на транспозицию равносильно выполнению этой транспозиции в нижней строке подстановки, т. е. меняет ее четность.

Удобным способом записи подстановок, позволяющим легко находить их четность, является *разложение в циклы.* Всякая подстановка n-й степени может некоторые из символов $1, 2, \ldots, n$ оставлять на месте, другие же действительно перемещать. *Циклической подстановкой* или *циклом* называется такая подстановка, что при повторении ее достаточное число раз всякий из действительно перемещаемых ею символов может быть переведен в любой другой из этих символов. Такова, например, подстановка восьмой степени

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 8 & 6 & 4 & 5 & 2 & 7 & 3 \end{pmatrix};$$

она действительно перемещает символы 2, 3, 6 и 8, причем переводит символ 2 в 8, символ 8 в 3, символ 3 в 6, а символ 6 снова в 2.

К числу циклов принадлежат все транспозиции. По аналогии с употребленной выше сокращенной записью транспозиций, для циклов употребляется следующая запись: действительно переставляемые символы записываются в круглых скобках друг за другом в том порядке, в каком они друг в друга переходят при повторении подстановки; начинается запись с любого из действительно перемещаемых символов, а последний символ считается переходящим в первый. Так, для указанного выше примера эта запись имеет вид

$$(2\ 8\ 3\ 6).$$

Число символов, действительно перемещаемых циклом, называется *длиной* цикла.

Два цикла n-й степени называются *независимыми*, если они не имеют общих действительно переставляемых символов. Понятно, что при перемножении независимых циклов порядок множителей не влияет на результат.

Всякая подстановка может быть единственным способом разложена в произведение попарно независимых циклов. Доказательство этого утверждения не представляет затруднений, и мы его опускаем. Практически разложение осуществляется следующим образом: начинаем с любого из действительно перемещаемых символов и выписываем за ним те символы, в которые он переходит при повторении подстановки, пока не вернемся к исходному символу. После этого «закрытия» цикла начинаем с одного из оставшихся действительно перемещаемых символов, получаем второй цикл и т. д.

Примеры.

1) $$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix} = (13)\,(254).$$

2) $$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 2 & 8 & 7 & 6 & 1 & 4 & 3 \end{pmatrix} = (156)\,(38)\,(47).$$

Обратно, для всякой подстановки, заданной разложением в независимые циклы, можно найти запись в обычной форме (при условии, что степень этой подстановки известна). Например,

3) $$(1372)\,(45) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 1 & 7 & 5 & 4 & 6 & 2 \end{pmatrix},$$

если известно, что степень этой подстановки есть 7.

Пусть дана подстановка n-й степени и пусть s есть число независимых циклов в ее разложении плюс число символов, оставляемых ею на месте[1]). Разность $n-s$ называется *декрементом* этой подстановки. Декремент равен, очевидно, числу действительно перемещаемых символов, уменьшенному на число независимых циклов, входящих в разложение подстановки. Для рассмотренных выше примеров 1), 2) и 3) декремент будет равен соответственно 3, 4 и 4.

Четность подстановки совпадает с четностью декремента этой подстановки.

Действительно, всякий цикл длины k можно следующим образом представить в виде произведения $k-1$ транспозиций:

$$(i_1,\ i_2,\ \ldots,\ i_k) = (i_1,\ i_2)\,(i_1,\ i_3)\ \ldots\ (i_1,\ i_k).$$

Предположим теперь, что дано разложение подстановки A в независимые циклы. Если каждый из циклов будет разложен указанным сейчас способом в произведение транспозиций, то мы получим представление подстановки A в виде произведения транспозиций. Число этих транспозиций будет, очевидно, меньше числа символов, действительно перемещаемых подстановкой A, на число, равное числу независимых циклов в разложении этой подстановки. Отсюда следует, что подстановку A можно разложить в произведение транспозиций, число которых равно декременту, а поэтому четность подстановки определяется четностью декремента.

---

[1]) Всякому символу, оставляемому подстановкой на месте, можно было бы поставить в соответствие «цикл» длины 1, т. е., например, в указанном выше примере 2) писать: (156) (38) (47) (2). Мы не будем, однако, этого делать.

§ 4. Определители n-го порядка

Мы хотим теперь обобщить результаты, полученные в § 2 для $n=2$ и 3, на случай произвольного n. Для этой цели необходимо ввести определители n-го порядка. Невозможно, однако, сделать это тем путем, каким были введены определители второго и третьего порядков, т. е. решая в общем виде системы линейных уравнений: по мере возрастания n вычисления становились бы все более и более громоздкими, а при произвольном n практически неосуществимыми. Мы выбираем иной путь: рассматривая уже известные нам определители второго и третьего порядков, мы постараемся установить общий закон, по которому эти определители выражаются через элементы соответствующих матриц, и применим этот закон в качестве определения для определителя порядка n, а затем докажем, что при таком определении правило Крамера остается справедливым.

Напомним выражения определителей второго и третьего порядков:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Мы видим, что всякий член определителя второго порядка есть произведение двух элементов, стоящих как в разных строках, так и в разных столбцах, причем все произведения такого вида, какие только можно составить из элементов матрицы второго порядка (их всего два), использованы в качестве членов определителя. Подобным же образом всякий член определителя третьего порядка является произведением трех элементов, также взятых по одному в каждой строке и в каждом столбце, причем снова все такие произведения используются в качестве членов определителя.

Пусть теперь дана квадратная матрица порядка n

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \tag{1}$$

Рассмотрим всевозможные произведения по n элементов этой матрицы, расположенных в разных строках и разных столбцах, т. е. произведения вида

$$a_{1\alpha_1}a_{2\alpha_2}\ \dots\ a_{n\alpha_n}, \tag{2}$$

где индексы $\alpha_1, \alpha_2, \dots, \alpha_n$ составляют некоторую перестановку из чисел $1, 2, \dots, n$. Число таких произведений равно числу различных

перестановок из n символов, т. е. равно $n!$. Будем считать все эти произведения членами будущего определителя n-го порядка, соответствующего матрице (1).

Для определения знака, с каким произведение (2) входит в состав определителя, заметим, что из индексов этого произведения можно составить подстановку

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}, \tag{3}$$

где i переходит в α_i, если в состав произведения (2) входит элемент, стоящий в i-й строке и α_i-м столбце матрицы (1). Рассматривая выражения определителей второго и третьего порядков, мы замечаем, что в них со знаком плюс входят те члены, индексы которых составляют четную подстановку, а со знаком минус — члены с нечетной подстановкой индексов. Естественно сохранить эту закономерность и в определении определителя n-го порядка.

Мы приходим, таким образом, к следующему определению: *определителем n-го порядка*, соответствующим матрице (1), называется алгебраическая сумма $n!$ членов, составленная следующим образом: членами служат всевозможные произведения n элементов матрицы, взятых по одному в каждой строке и в каждом столбце, причем член берется со знаком плюс, если его индексы составляют четную подстановку, и со знаком минус — в противоположном случае.

Для записи определителя n-го порядка, соответствующего матрице (1), мы будем, как и в случае определителей второго и третьего порядков, употреблять символ

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{4}$$

Определители n-го порядка превращаются при $n=2$ и $n=3$ в рассмотренные ранее определители второго и третьего порядков, а при $n=1$, т. е. для матриц, состоящих из одного элемента, определитель равен самому этому элементу. Мы не знаем пока, однако, можно ли при $n>3$ использовать определитель n-го порядка для решения систем линейных уравнений. Это будет показано в § 7; предварительно необходимо подвергнуть определители n-го порядка детальному изучению и, в частности, найти методы для их вычисления, так как вычислять определители, непосредственно применяя их определение, даже при не очень больших n было бы весьма затруднительным.

Сейчас мы установим некоторые простейшие свойства определителей n-го порядка, относящиеся преимущественно к одному из следующих двух вопросов: с одной стороны, нас будут интересовать условия, при которых определитель равен нулю; с другой стороны,

мы укажем некоторые преобразования матрицы, которые не меняют ее определителя или же подвергают его легко учитываемым изменениям.

Назовем *транспонированием* матрицы (1) такое преобразование этой матрицы, при котором ее строки делаются столбцами с тем же самым номером, т. е. переход от матрицы (1) к матрице

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}; \tag{5}$$

можно сказать, что транспонирование есть поворот матрицы (1) около главной диагонали. Соответственно говорят, что определитель

$$\begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} \tag{6}$$

получен транспонированием определителя (4).

Свойство 1. *Определитель не меняется при транспонировании.*

В самом деле, всякий член определителя (4) имеет вид

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}, \tag{7}$$

где вторые индексы составляют некоторую перестановку из символов $1, 2, \dots, n$. Однако все множители произведения (7) и в определителе (6) остаются в разных строках и разных столбцах, т. е. (7) служит членом и для транспонированного определителя. Верно, очевидно, и обратное, и поэтому определители (4) и (6) состоят из одних и тех же членов. Знак члена (7) в определителе (4) определяется четностью подстановки

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}; \tag{8}$$

в определителе (6) первые индексы элементов указывают на номер столбца, вторые — на номер строки, поэтому члену (7) в определителе (6) соответствует подстановка

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 2 & \dots & n \end{pmatrix}. \tag{9}$$

Подстановки (8) и (9) — в общем случае различные, но имеют, очевидно, одну и ту же четность, а поэтому член (7) имеет в обоих определителях один и тот же знак. Таким образом, определители (4) и (6) являются суммами одинаковых членов, взятых с одинаковыми знаками, т. е. равны друг другу.

Из свойства 1 вытекает, что всякое утверждение о строках определителя справедливо и для его столбцов и обратно, т. е. что *в определителе* (в отличие от матрицы) *строки и столбцы равноправны*. Исходя из этого, мы будем дальнейшие свойства 2—9 формулировать и доказывать лишь для строк определителя; аналогичные свойства для столбцов не будут требовать особого доказательства.

Свойство 2. *Если одна из строк определителя состоит из нулей, то определитель равен нулю.*

Действительно, пусть все элементы i-й строки определителя являются нулями. В каждый член определителя должен войти множителем один элемент из i-й строки, поэтому в нашем случае все члены определителя равны нулю.

Свойство 3. *Если один определитель получен из другого перестановкой двух строк, то все члены первого определителя будут членами и во втором, но с обратными знаками, т. е. от перестановки двух строк определитель лишь меняет знак.*

В самом деле, пусть в определителе (4) переставляются i-я и j-я строки, $i \neq j$, а все остальные строки остаются на месте. Мы получаем определитель

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \begin{matrix} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{matrix} \qquad (10)$$

(сбоку указаны номера строк). Если

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} \qquad (11)$$

есть член определителя (4), то все его множители и в определителе (10) остаются, очевидно, в разных строках и разных столбцах. Таким образом, определители (4) и (10) состоят из одних и тех же членов. Члену (11) в определителе (4) соответствует подстановка

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (12)$$

а в определителе (10) — подстановка

$$\begin{pmatrix} 1 & 2 & \dots & j & \dots & i & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (13)$$

так как, например, элемент $a_{i\alpha_i}$ стоит теперь в j-й строке, но остается в старом α_i-м столбце. Подстановка (13) получается, однако, из подстановки (12) путем одной транспозиции в верхней строке, т. е. имеет противоположную четность. Отсюда следует, что все члены определителя (4) входят в определитель (10) с обратными знаками, т. е. определители (4) и (10) отличаются друг от друга лишь знаком.

Свойство 4. *Определитель, содержащий две одинаковые строки, равен нулю.*

В самом деле, пусть определитель равен числу d и пусть соответственные элементы его i-й и j-й строк $(i \neq j)$ равны между собой. После перестановки этих двух строк определитель станет равен, ввиду свойства 3, числу $-d$. Так как, однако, переставляются одинаковые строки, то определитель на самом деле не меняется, т. е. $d = -d$, откуда $d = 0$.

Свойство 5. *Если все элементы некоторой строки определителя умножить на некоторое число k, то сам определитель умножится на k.*

Пусть на k умножены все элементы i-й строки. Каждый член определителя содержит ровно один элемент из i-й строки, поэтому всякий член приобретает множитель k, т. е. сам определитель умножается на k.

Это свойство допускает и такую формулировку: *общий множитель всех элементов некоторой строки определителя можно вынести за знак определителя.*

Свойство 6. *Определитель, содержащий две пропорциональные строки, равен нулю.*

В самом деле, пусть элементы j-й строки определителя отличаются от соответствующих элементов i-й строки $(i \neq j)$ одним и тем же множителем k. Вынося этот общий множитель k из j-й строки за знак определителя, мы получим определитель с двумя одинаковыми строками, равный нулю по свойству 4.

Свойство 4 (а также свойство 2 при $n > 1$) является, очевидно, частными случаями свойства 6 (при $k = 1$ и $k = 0$).

Свойство 7. *Если все элементы i-й строки определителя n-го порядка представлены в виде суммы двух слагаемых:*

$$a_{ij} = b_j + c_j, \quad j = 1, \ldots, n,$$

то определитель равен сумме двух определителей, у которых все строки, кроме i-й, — такие же, как и в заданном определителе, а i-я строка в одном из слагаемых состоит из элементов b_j, в другом — из элементов c_j.

Действительно, всякий член заданного определителя можно представить в виде

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{i\alpha_i} \ldots a_{n\alpha_n} = a_{1\alpha_1} a_{2\alpha_2} \ldots (b_{\alpha_i} + c_{\alpha_i}) \ldots a_{n\alpha_n} =$$
$$= a_{1\alpha_1} a_{2\alpha_2} \ldots b_{\alpha_i} \ldots a_{n\alpha_n} + a_{1\alpha_1} a_{2\alpha_2} \ldots c_{\alpha_i} \ldots a_{n\alpha_n}.$$

Собирая вместе первые слагаемые этих сумм (с теми же знаками, какие имели соответствующие члены в заданном определителе), мы получим, очевидно, определитель n-го порядка, отличающийся от заданного определителя лишь тем, что в i-й строке вместо элементов a_{ij} стоят элементы b_j. Соответственно вторые слагаемые составляют

определитель, в i-й строке которого стоят элементы c_j. Таким образом,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ b_1+c_1 & b_2+c_2 & \dots & b_n+c_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ b_1 & b_2 & \dots & b_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ c_1 & c_2 & \dots & c_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Свойство 7 без труда распространяется на случай, когда всякий элемент i-й строки есть сумма не двух, а m слагаемых, $m \geqslant 2$.

Будем говорить, что i-я строка определителя есть *линейная комбинация* его остальных строк, если для всякой строки с номером j, $j = 1, \dots, i-1, i+1, \dots, n$, можно указать такое число k_j, что, умножая j-ю строку на k_j, а затем складывая все строки, кроме i-й (причем сложение строк следует понимать так, что складываются элементы всех этих строк в каждом столбце отдельно), мы получим i-ю строку. Некоторые из коэффициентов k_j могут быть равными нулю, т. е. i-я строка будет на самом деле линейной комбинацией не всех, а лишь некоторых из оставшихся строк. В частности, если лишь один из коэффициентов k_j отличен от нуля, мы получаем случай пропорциональности двух строк. Наконец, если строка состоит целиком из нулей, то она всегда будет линейной комбинацией остальных строк, — случай, когда все k_j равны нулю.

Свойство 8. *Если одна из строк определителя есть линейная комбинация его других строк, то определитель равен нулю.*

Пусть, например, i-я строка будет линейной комбинацией s других строк, $1 \leqslant s \leqslant n-1$. Всякий элемент i-й строки будет тогда суммой s слагаемых, а поэтому, применяя свойство 7, мы представим наш определитель в виде суммы определителей, в каждом из которых i-я строка будет пропорциональна одной из других строк. По свойству 6 все эти определители равны нулю; равен нулю, следовательно, и заданный определитель.

Это свойство является обобщением свойства 6, причем, как будет доказано в § 10, оно дает самый общий случай равенства определителя нулю.

Свойство 9. *Определитель не меняется, если к элементам одной из его строк прибавляются соответственные элементы другой строки, умноженные на одно и то же число.*

Пусть, в самом деле, к i-й строке определителя d прибавляется j-я строка, $j \neq i$, умноженная на число k, т. е. в новом определителе всякий элемент i-й строки имеет вид $a_{is} + ka_{js}$, $s = 1, 2, \dots, n$. Тогда, на основании свойства 7, этот определитель равен сумме двух определителей, из которых первый есть d, а второй содержит две пропорциональные строки и поэтому равен нулю.

Так как число k может быть и отрицательным, то *определитель не меняется и при вычитании из одной его строки другой строки,*

умноженной на некоторое число. Вообще, *определитель не меняется, если к одной из его строк прибавляется любая линейная комбинация других строк.*

Рассмотрим один пример. Определитель называется *кососимметрическим*, если его элементы, симметричные относительно главной диагонали, отличаются друг от друга лишь знаком, т. е. если при всех i и j будет $a_{ji} = -a_{ij}$; отсюда следует, что для всех i будет $a_{ii} = -a_{ii} = 0$. Таким образом, определитель имеет вид

$$d = \begin{vmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ -a_{12} & 0 & a_{23} & \dots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ -a_{1n} & -a_{2n} & -a_{3n} & \dots & 0 \end{vmatrix}.$$

Умножая каждую строку этого определителя на число -1, мы получим транспонированный определитель, т. е. снова равный d, откуда, ввиду свойства 5, следует:

$$(-1)^n d = d.$$

При нечетном n отсюда вытекает: $-d = d$, т. е. $d = 0$. Таким образом, *всякий кососимметрический определитель нечетного порядка равен нулю.*

§ 5. Миноры и их алгебраические дополнения

Выше уже отмечалось, что было бы затруднительно вычислять определители n-го порядка, применяя непосредственно их определение, т. е. каждый раз выписывая все $n!$ членов, определяя их знаки и т. д. Существуют более простые методы вычисления определителей, основанные на том, что определитель порядка n может быть выражен через определители более низких порядков. С этой целью введем следующее понятие.

Пусть дан определитель d порядка n. Берем целое число k, удовлетворяющее условию $1 \leqslant k \leqslant n-1$, и в определителе d выбираем произвольные k строк и k столбцов. Элементы, стоящие на пересечении этих строк и столбцов, т. е. принадлежащие к одной из выбранных строк и к одному из выбранных столбцов, составляют, очевидно, матрицу порядка k. Определитель этой матрицы называется *минором k-го порядка* определителя d. Можно сказать также, что минор k-го порядка есть определитель, получающийся после вычеркивания в определителе d $n-k$ строк и $n-k$ столбцов. В частности, после вычеркивания в определителе одной строки и одного столбца мы получаем минор $(n-1)$-го порядка; с другой стороны, минорами первого порядка будут отдельные элементы определителя d.

Пусть в определителе d n-го порядка взят минор M k-го порядка. Если мы вычеркнем те строки и столбцы, на пересечении которых стоит этот минор, то останется минор M' $(n-k)$-го порядка, который называется *дополнительным минором* для минора M. Если мы вычеркнем, наоборот, те строки и столбцы, в которых расположены

элементы минора M', то останется, очевидно, минор M. Таким образом, можно говорить о паре взаимно дополнительных миноров определителя. В частности, элемент a_{ij} и минор $(n-1)$-го порядка, получающийся вычеркиванием в определителе i-й строки и j-го столбца, будут составлять пару взаимно дополнительных миноров.

Если минор k-го порядка M расположен в строках с номерами $i_1, i_2, \ldots, i_k$ и в столбцах с номерами $j_1, j_2, \ldots, j_k$, то назовем *алгебраическим дополнением* минора M его дополнительный минор M', взятый со знаком плюс или минус в зависимости от того, четна или нечетна сумма номеров всех строк и столбцов, в которых расположен минор M, т. е. сумма

$$s_M = i_1 + i_2 + \ldots + i_k + j_1 + j_2 + \ldots + j_k. \tag{1}$$

Иными словами, алгебраическим дополнением для минора M будет число $(-1)^{s_M} M'$.

Произведение любого минора M k-го порядка на его алгебраическое дополнение в определителе d является алгебраической суммой, слагаемые которой, получающиеся от умножения членов минора M на взятые со знаком $(-1)^{s_M}$ члены дополнительного минора M', будут некоторыми членами определителя d, причем их знаки в этой сумме совпадают с теми знаками, с какими они входят в состав определителя.

Доказательство этой теоремы мы начнем со случая, когда минор M расположен в левом верхнем углу определителя:

$$d = \left|\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & M & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ \hline a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & M' & \ldots \\ a_{n1} & \ldots & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{array}\right|,$$

т. е. в строках с номерами $1, 2, \ldots, k$ и в столбцах с такими же номерами. Тогда минор M' будет занимать правый нижний угол определителя. Число s_M в этом случае будет четным:

$$s_M = 1 + 2 + \ldots + k + 1 + 2 + \ldots + k = 2(1 + 2 + \ldots + k),$$

поэтому алгебраическим дополнением для M служит сам минор M'.

Берем произвольный член

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} \tag{2}$$

минора M; его знак в M будет $(-1)^l$, если l есть число инверсий в подстановке

$$\begin{pmatrix} 1 & 2 & \ldots & k \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k \end{pmatrix}. \tag{3}$$

Произвольный член

$$a_{k+1,\ \beta_{k+1}} a_{k+2,\ \beta_{k+2}} \ldots a_{n\beta_n} \tag{4}$$

минора M' имеет в этом миноре знак $(-1)^{l'}$, где l' есть число инверсий в подстановке

$$\begin{pmatrix} k+1 & k+2 & \ldots & n \\ \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix}. \tag{5}$$

Перемножая члены (2) и (4), мы получим произведение n элементов

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} a_{k+1,\ \beta_{k+1}} a_{k+2,\ \beta_{k+2}} \ldots a_{n\beta_n}, \tag{6}$$

расположенных в разных строках и разных столбцах определителя; оно будет, следовательно, членом определителя d. Знак члена (6) в произведении MM' будет произведением знаков членов (2) и (4), т. е. $(-1)^l \cdot (-1)^{l'} = (-1)^{l+l'}$. Такой же знак имеет, однако, член (6) и в определителе d. Действительно, нижняя строка подстановки

$$\begin{pmatrix} 1 & 2 & \ldots & k & k+1 & k+2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k & \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix},$$

составленной из индексов этого члена, содержит лишь $l+l'$ инверсий, так как никакое α ни с одним из β не может составить инверсию: все α не больше k, все β не меньше $k+1$.

Этим доказан рассматриваемый нами частный случай теоремы. Переходим к рассмотрению общего случая, т. е. предположим, что минор M расположен в строках с номерами $i_1, i_2, \ldots, i_k$ и в столбцах с номерами $j_1, j_2, \ldots, j_k$, причем

$$i_1 < i_2 < \ldots < i_k, \quad j_1 < j_2 < \ldots < j_k.$$

Постараемся, переставляя строки и столбцы определителя, передвинуть минор M в левый верхний угол, причем так, чтобы дополнительный минор не изменился. Для этой цели переставим i_1-ю строку с (i_1-1)-й, затем с (i_1-2)-й и т. д., пока i_1-я строка не займет место первой строки; для этого мы должны будем переставить строки i_1-1 раз. Будем затем последовательно переставлять i_2-ю строку со строками, расположенными над нею, пока она не расположится непосредственно под i_1-й строкой, т. е. на месте, которое до начала всех преобразований занимала вторая строка; для этого, как легко проверить, мы должны будем переставить строки i_2-2 раза. Аналогичным образом i_3-ю строку мы передвинем на место третьей строки и т. д., пока i_k-я строка не окажется на месте k-й строки. Всего мы должны будем совершить

$$(i_1-1)+(i_2-2)+\ldots+(i_k-k)=$$
$$=(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)$$

транспозиций строк.

Минор M расположен уже в первых k строках нового определителя. Будем теперь последовательно переставлять столбцы определителя: j_1-й со всеми ему предшествующими, пока он не займет первого места, затем j_2-й, пока он не займет второго места, и т. д. Всего столбцы будут переставлены

$$(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)$$

раз.

После всех этих преобразований мы приходим к новому определителю d', в котором минор M занимает левый верхний угол. Так как мы переставляли каждый раз лишь соседние строки или столбцы, то взаимное расположение строк и столбцов, содержавших в определителе d минор M', остается без изменения, а поэтому дополнительным к минору M в определителе d' остается минор M', занимающий, однако, уже правый нижний угол. Как доказано выше, произведение MM' является суммой некоторого количества членов определителя d', взятых с теми же знаками, с какими они входят в d'. Однако определитель d' получен из определителя d путем

$$[(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)]+\\+[(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)]=\\=s_M-2(1+2+\ldots+k)$$

транспозиций строк и столбцов, поэтому, как мы знаем из предшествующего параграфа, члены определителя d' отличаются от соответственных членов определителя d лишь знаком $(-1)^{s_M}$ (четное число $2(1+2+\ldots+k)$ не будет, понятно, влиять на знак). Отсюда следует, что произведение $(-1)^{s_M}MM'$ состоит из некоторого количества членов определителя d, взятых с такими же знаками, какие они имеют в этом определителе. Теорема доказана.

Заметим, что если миноры M и M' взаимно дополнительны, то числа s_M и $s_{M'}$ имеют одинаковую четность. Действительно, номер всякой строки и всякого столбца входит слагаемым в одно и только одно из этих чисел, а поэтому сумма $s_M+s_{M'}$ равна общей сумме номеров всех строк и столбцов определителя, т. е. равна четному числу $2(1+2+\ldots+n)$.

§ 6. Вычисление определителей

Результаты предшествующего параграфа позволяют свести вычисление определителя n-го порядка на вычисление нескольких определителей $(n-1)$-го порядка. Введем сначала следующие обозначения: если a_{ij} — элемент определителя d, то через M_{ij} обозначим дополнительный минор или, короче, *минор этого элемента*, т. е. минор $(n-1)$-го порядка, получающийся после вычеркивания из определи-

теля i-й строки и j-го столбца. Далее, через A_{ij} обозначим алгебраическое дополнение элемента a_{ij}, т. е.

$$A_{ij}=(-1)^{i+j}M_{ij}.$$

Как доказано в предшествующем параграфе, произведение $a_{ij}A_{ij}$ является суммой нескольких членов определителя d, входящих в эту сумму с теми же знаками, с какими они входят в состав определителя d. Легко подсчитать число этих членов: оно равно числу членов в миноре M_{ij}, т. е. равно $(n-1)!$.

Выбираем теперь любую i-ю строку определителя d и берем произведение каждого элемента этой строки на его алгебраическое дополнение:

$$a_{i1}A_{i1},\ a_{i2}A_{i2},\ \ldots,\ a_{in}A_{in}. \qquad (1)$$

Никакой член определителя d не может войти в состав двух разных из числа произведений (1): все члены определителя, входящие в произведение $a_{i1}A_{i1}$, содержат из i-й строки элемент a_{i1} и поэтому отличаются от членов, входящих в произведение $a_{i2}A_{i2}$, т. е. содержащих из i-й строки элемент a_{i2}, и т. д.

С другой стороны, общее число членов определителя d, входящих во все произведения (1), равно

$$(n-1)!\cdot n=n!,$$

т. е. этим исчерпываются вообще все члены определителя d. Мы доказали, таким образом, что имеет место следующее *разложение определителя d по i-й строке*:

$$d=a_{i1}A_{i1}+a_{i2}A_{i2}+\ldots+a_{in}A_{in}, \qquad (2)$$

т. е. *определитель d равен сумме произведений всех элементов произвольной его строки на их алгебраические дополнения.* Аналогичное разложение определителя можно получить и по любому его столбцу.

Заменяя в разложении (2) алгебраические дополнения соответствующими минорами со знаками плюс или минус, мы сведем вычисление определителя n-го порядка к вычислению нескольких определителей $(n-1)$-го порядка. Заметим, что если некоторые из элементов i-й строки равны нулю, то соответствующие им миноры не нужно будет, понятно, вычислять. Ввиду этого полезно предварительно так преобразовать определитель, используя свойство 9 (см. § 4), чтобы в одной из строк или в одном из столбцов достаточно много элементов оказалось замененными нулями. В действительности *свойство 9 позволяет в любой строке или любом столбце заменить нулями все элементы, кроме одного.* В самом деле, если $a_{ik}\neq 0$, то любой элемент i-й строки a_{ij}, $j\neq k$, будет заменен нулем после вычитания k-го столбца, умноженного

на $\frac{a_{ij}}{a_{ik}}$, из j-го столбца. Таким образом, вычисление определителя n-го порядка можно свести к вычислению **одного** определителя $(n-1)$-го порядка.

Примеры.

1. Вычислить определитель четвертого порядка

$$d=\begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}.$$

Разложим его по третьей строке, используя наличие в ней одного нуля:

$$d=(-1)^{3+1}\cdot 2\cdot\begin{vmatrix} 1 & -1 & 2 \\ 1 & 3 & -4 \\ -5 & 3 & -3 \end{vmatrix}+$$

$$+(-1)^{3+3}\cdot 1\cdot\begin{vmatrix} 3 & 1 & 2 \\ -5 & 1 & -4 \\ 1 & -5 & -3 \end{vmatrix}+(-1)^{3+4}\cdot(-1)\cdot\begin{vmatrix} 3 & 1 & -1 \\ -5 & 1 & 3 \\ 1 & -5 & 3 \end{vmatrix}.$$

Вычисляя полученные определители третьего порядка, получим:

$$d=2\cdot 16-40+48=40.$$

2. Вычислить определитель пятого порядка

$$d=\begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & 0 & 3 & 7 & -2 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 6 & -4 & 1 & 2 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

Прибавляя ко второй строке утроенную пятую и вычитая из четвертой строки учетверенную пятую, получим:

$$d=\begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & -9 & 0 & 13 & 7 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 18 & 0 & -7 & -10 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

Разложив этот определитель по третьему столбцу, содержащему лишь один не равный нулю элемент (с суммой индексов $5+3$, т. е. четной), получим

$$d=(-1)\cdot\begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix}.$$

Преобразуем вновь полученный определитель, прибавляя к первой строке удвоенную вторую и вычитая из третьей строки утроенную вторую, а из четвертой — удвоенную вторую:

$$d=-\begin{vmatrix} 0 & -13 & 25 & 17 \\ 1 & -9 & 13 & 7 \\ 0 & 26 & -34 & -26 \\ 0 & 36 & -33 & -24 \end{vmatrix},$$

а затем разложим его по первому столбцу, причем заметим, что единственному не равному нулю элементу этого столбца соответствует нечетная сумма индексов, получим:

$$d=\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix}.$$

Вычислим этот определитель третьего порядка, предварительно разложив его по третьей строке:

$$d=36\cdot\begin{vmatrix} 25 & 17 \\ -34 & -26 \end{vmatrix}-(-33)\cdot\begin{vmatrix} -13 & 17 \\ 26 & -26 \end{vmatrix}+(-24)\cdot\begin{vmatrix} -13 & 25 \\ 26 & -34 \end{vmatrix}=$$

$$=36\cdot(-72)-(-33)\cdot(-104)+(-24)\cdot(-208)=-1032.$$

3. *Если все элементы определителя, расположенные по одну сторону от главной диагонали, равны нулю, то этот определитель равен произведению элементов, стоящих на главной диагонали.*

Для определителя второго порядка это утверждение очевидно. Мы будем поэтому доказывать его по индукции, т. е. предположим, что для определителей $(n-1)$-го порядка оно уже доказано, и рассмотрим определитель n-го порядка

$$d=\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Разлагая его по первому столбцу, получим:

$$d=a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Но к минору, стоящему в правой части, применимо предположение индукции, т. е. он равен произведению $a_{22}a_{33}\dots a_{nn}$, а поэтому

$$d=a_{11}a_{22}a_{33}\dots a_{nn}.$$

4. *Определителем Вандермонда* называется определитель

$$d=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}.$$

Докажем, что *при любом n определитель Вандермонда равен произведению всевозможных разностей a_i-a_j, где $1 \leqslant j < i \leqslant n$.* Действительно при $n=2$ будет

$$\begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2-a_1.$$

Пусть наше утверждение уже доказано для определителей Вандермонда $(n-1)$-го порядка. Преобразуем определитель d следующим образом: из n-й (последней) строки вычитаем $(n-1)$-ю, умноженную на a_1, затем из $(n-1)$-й вычитаем $(n-2)$-ю, также умноженную на a_1, и т. д., наконец, из второй строки вычитаем первую, умноженную на a_1. Мы получим:

$$d=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & a_2-a_1 & a_3-a_1 & \dots & a_n-a_1 \\ 0 & a_2^2-a_1a_2 & a_3^2-a_1a_3 & \dots & a_n^2-a_1a_n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_2^{n-1}-a_1a_2^{n-2} & a_3^{n-1}-a_1a_3^{n-2} & \dots & a_n^{n-1}-a_1a_n^{n-2} \end{vmatrix}.$$

Разлагая этот определитель по первому столбцу, мы придем к определителю $(n-1)$-го порядка; после вынесения из всех столбцов общих множителей за знак определителя он примет вид

$$d=(a_2-a_1)(a_3-a_1)\dots(a_n-a_1)\cdot\begin{vmatrix} 1 & 1 & \dots & 1 \\ a_2 & a_3 & \dots & a_n \\ a_2^2 & a_3^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_2^{n-2} & a_3^{n-2} & \dots & a_n^{n-2} \end{vmatrix}.$$

Последний множитель является определителем Вандермонда $(n-1)$-го порядка, т. е., по предположению, равен произведению всех разностей a_i-a_j для $2 \leqslant j < i \leqslant n$. Можно написать, следовательно, употребляя символ $\prod$ для обозначения произведения, что

$$d=(a_2-a_1)(a_3-a_1)\dots(a_n-a_1)\cdot\prod_{2\leqslant j<i\leqslant n}(a_i-a_j)=\prod_{1\leqslant j<i\leqslant n}(a_i-a_j).$$

Таким же методом может быть доказано, что определитель

$$d'=\begin{vmatrix} a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ a_1 & a_2 & a_3 & \dots & a_n \\ 1 & 1 & 1 & \dots & 1 \end{vmatrix}$$

равен произведению всевозможных разностей $a_i - a_j$, где $1 \leqslant i < j \leqslant n$, т. е.

$$d' = \prod_{1 \leqslant i < j \leqslant n} (a_i - a_j).$$

Обобщая полученные выше разложения определителя по строке или столбцу, докажем следующую теорему, говорящую о *разложении определителя по нескольким строкам или столбцам.*

Теорема Лапласа. *Пусть в определителе d порядка n произвольно выбраны k строк (или k столбцов), $1 \leqslant k \leqslant n-1$. Тогда сумма произведений всех миноров k-го порядка, содержащихся в выбранных строках, на их алгебраические дополнения равна определителю d.*

Доказательство. Пусть в определителе d выбраны строки с номерами $i_1, i_2, \ldots, i_k$. Мы знаем, что произведение любого минора k-го порядка M, расположенного в этих строках, на его алгебраическое дополнение состоит из некоторого количества членов определителя d, взятых с теми же знаками, с какими они входят в состав определителя. Теорема будет, следовательно, доказана, если мы покажем, что, заставляя M пробегать все миноры k-го порядка, расположенные в выбранных строках, мы получим все члены определителя, причем ни один из них не встретится дважды.

Пусть

$$a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n} \tag{3}$$

— произвольный член определителя d. Возьмем отдельно произведение тех элементов из этого члена, которые принадлежат к выбранным нами строкам с номерами $i_1, i_2, \ldots, i_k$. Это будет произведение

$$a_{i_1\alpha_{i_1}} a_{i_2\alpha_{i_2}} \ldots a_{i_k\alpha_{i_k}}; \tag{4}$$

k множителей этого произведения стоят в k различных столбцах, а именно, в столбцах с номерами $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$. Эти номера столбцов вполне определяются, следовательно, заданием члена (3). Если мы обозначим через M минор k-го порядка, стоящий на пересечении столбцов с этими номерами $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$ и выбранных ранее строк с номерами $i_1, i_2, \ldots, i_k$, то произведение (4) будет одним из членов минора M, а произведение всех элементов из члена (3), не вошедших в (4), — членом его дополнительного минора. Таким образом, всякий член определителя входит в произведение некоторого, притом вполне определенного, минора k-го порядка из выбранных строк на его дополнительный минор, причем является произведением вполне определенных членов этих двух миноров. Для того же, наконец, чтобы получить взятый нами член определителя с тем знаком, какой он имеет в определителе, остается, как мы знаем, заменить дополнительный минор алгебраическим дополнением. Этим заканчивается доказательство теоремы.

Доказательство теоремы можно было бы вести и несколько иным путем. Именно, произведение любого минора k-го порядка M, расположенного в выбранных строках, на его алгебраическое дополнение состоит из $k!\,(n-k)!$ членов, так как минор k-го порядка M состоит из $k!$ членов, а его алгебраическое дополнение, отличаясь, возможно, лишь знаком от минора порядка $n-k$, содержит $(n-k)!$ членов. С другой стороны, число миноров k-го порядка, содержащихся в выбранных нами строках, равно числу сочетаний из n по k, т. е. равно числу

$$\frac{n!}{k!\,(n-k)!}$$

Перемножая, мы получаем, что сумма произведений всех миноров k-го порядка из выбранных строк на их алгебраические дополнения состоит из $n!$ слагаемых. Таково же, однако, и общее число членов определителя d. Теорема будет, следовательно, доказана, если мы покажем, что всякий член определителя d входит хотя бы один раз (а тогда и точно один раз) в рассматриваемую сумму произведений миноров на их алгебраические дополнения. Для этого читателю остается повторить (с некоторыми упрощениями) рассуждения, проведенные в предшествующем доказательстве.

Теорема Лапласа позволяет сводить вычисление определителя n-го порядка к вычислению нескольких определителей порядков k и $n-k$. Этих новых определителей окажется, вообще говоря, весьма много, и поэтому применять теорему Лапласа целесообразно лишь в том случае, если в определителе можно так выбрать k строк (или столбцов), что многие из миноров k-го порядка, расположенных в этих строках, будут равны нулю.

Примеры.

1. Пусть дан определитель, все элементы которого, стоящие в первых k строках и последних $n-k$ столбцах, равны нулю:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} & & & \\ \cdot & \cdot & \cdot & & 0 & \\ a_{k1} & \dots & a_{kk} & & & \\ a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nk} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Тогда этот определитель равен произведению двух своих миноров:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} \end{vmatrix}\cdot\begin{vmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot \\ a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Для доказательства достаточно разложить определитель по первым k строкам.

2. Пусть дан определитель d порядка $2n$, в левом верхнем углу которого стоит минор n-го порядка, составленный целиком из нулей. Если миноры n-го порядка, стоящие в правом верхнем, левом нижнем и правом нижнем

углах определителя, будут обозначены соответственно через M, M' и M'', т. е. определитель можно символически записать в виде $d=\begin{vmatrix} 0 & M \\ M' & M'' \end{vmatrix}$, то $d=(-1)^n MM'$.

Для доказательства разложим определитель по первым n строкам и заметим, что

$$s_M=(1+2+\ldots+n)+[(n+1)+(n+2)\ldots+2n]=n+2n^2,$$

т. е. s_M и n имеют одинаковую четность.

3. Вычислить определитель

$$d=\begin{vmatrix} -4 & 1 & 2 & -2 & 1 \\ 0 & 3 & 0 & 1 & -5 \\ 2 & -3 & 1 & -3 & 1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 4 & 0 & 2 & 5 \end{vmatrix}.$$

Разлагая его по первому и третьему столбцам, содержащим удачно расположенные нули, мы получим:

$$d=(-1)^{1+3+1+3}\begin{vmatrix} -4 & 2 \\ 2 & 1 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix}+$$

$$+(-1)^{1+4+1+3}\begin{vmatrix} -4 & 2 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix}+$$

$$+(-1)^{3+4+1+3}\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ 4 & 2 & 5 \end{vmatrix}=$$

$$=(-8)\cdot(-20)-(-10)\cdot(-62)-7\cdot 87=-1069.$$

§ 7. Правило Крамера

Изложенная выше теория определителей n-го порядка позволяет показать, что эти определители, введенные лишь по аналогии с определителями второго и третьего порядков, подобно последним могут быть использованы для решения систем линейных уравнений. Сначала сделаем, впрочем, одно дополнительное замечание, связанное с разложениями определителей по строке или столбцу; это замечание будет в дальнейшем неоднократно использоваться.

Разложим определитель

$$d=\begin{vmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2j} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{vmatrix}.$$

по его j-му столбцу:

$$d = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj},$$

а затем заменим в этом разложении элементы j-го столбца системой n произвольных чисел $b_1, b_2, \ldots, b_n$. Выражение

$$b_1A_{1j} + b_2A_{2j} + \ldots + b_nA_{nj},$$

которое мы получим, будет, очевидно, служить разложением по j-му столбцу для определителя

$$d' = \begin{vmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ a_{21} & \ldots & b_2 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{vmatrix},$$

получающегося из определителя d заменой его j-го столбца столбцом из чисел $b_1, b_2, \ldots, b_n$. В самом деле, замена j-го столбца определителя d не влияет на миноры элементов этого столбца, а поэтому и на их алгебраические дополнения.

Применим это к случаю, когда в качестве чисел $b_1, b_2, \ldots, b_n$ берутся элементы k-го столбца определителя d при $k \neq j$. Определитель, который мы получим после такой замены, будет содержать два одинаковых столбца (j-й и k-й) и поэтому будет равен нулю. Равно нулю, следовательно, и разложение этого определителя по его j-му столбцу, т. е.

$$a_{1k}A_{1j} + a_{2k}A_{2j} + \ldots + a_{nk}A_{nj} = 0 \text{ при } j \neq k.$$

Таким образом, *сумма произведений всех элементов некоторого столбца определителя на алгебраические дополнения соответственных элементов другого столбца равна нулю.* Такой же результат справедлив, конечно, и для строк определителя.

Переходим к рассмотрению систем линейных уравнений, причем ограничимся пока случаем систем, в которых число уравнений равно числу неизвестных, т. е. систем вида

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n. \end{aligned}\right\} \quad (1)$$

Дополнительно предположим, что определитель d из коэффициентов при неизвестных в этой системе, называемый кратко *определителем системы*, отличен от нуля. При этих предположениях мы докажем, что система (1) совместна и даже определенна.

В § 2, решая систему трех уравнений с тремя неизвестными, мы умножали каждое из уравнений на некоторый множитель, а затем складывали эти уравнения, после чего коэффициенты при двух неизвестных из трех оказывались равными нулю. Сейчас мы

легко обнаруживаем, что множители, которые нами употреблялись, были алгебраическими дополнениями в определителе системы к элементу, являющемуся коэффициентом при искомом неизвестном в данном уравнении. Этот же прием будет теперь использован для решения системы (1).

Предположим сначала, что система (1) совместна и $\alpha_1, \alpha_2, \ldots, \alpha_n$ — одно из ее решений. Справедливы, следовательно, равенства

$$\left.\begin{array}{l} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = b_1, \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\alpha_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \ldots + a_{nn}\alpha_n = b_n. \end{array}\right\} \quad (2)$$

Пусть j будет любым из чисел $1, 2, \ldots, n$. Умножим обе части первого из равенств (2) на A_{1j}, т. е. на алгебраическое дополнение элемента a_{1j} в определителе системы d; обе части второго равенства умножим на A_{2j} и т. д., наконец, обе части последнего — на A_{nj}. Складывая затем отдельно левые и отдельно правые части всех равенств, мы придем к следующему равенству:

$$\begin{aligned} (a_{11}A_{1j} + a_{21}A_{2j} + \ldots + a_{n1}A_{nj})\,\alpha_1 + \\ + (a_{12}A_{1j} + a_{22}A_{2j} + \ldots + a_{n2}A_{nj})\,\alpha_2 + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + (a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj})\,\alpha_j + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + (a_{1n}A_{1j} + a_{2n}A_{2j} + \ldots + a_{nn}A_{nj})\,\alpha_n = \\ = b_1A_{1j} + b_2A_{2j} + \ldots + b_nA_{nj}. \end{aligned}$$

Коэффициентом при α_j в этом равенстве служит d, коэффициенты при всех остальных α будут, ввиду сделанного выше замечания, равны нулю, а свободный член будет определителем, получающимся из определителя d после замены в нем j-го столбца столбцом из свободных членов системы (1). Если этот последний определитель мы обозначим, как и в § 2, через d_j, то наше равенство примет вид

$$d\alpha_j = d_j,$$

откуда, ввиду $d \neq 0$,

$$\alpha_j = \frac{d_j}{d}.$$

Этим доказано, что если система (1) совместна, то она обладает единственным решением

$$\alpha_1 = \frac{d_1}{d}, \quad \alpha_2 = \frac{d_2}{d}, \quad \ldots, \quad \alpha_n = \frac{d_n}{d}. \quad (3)$$

Покажем теперь, что система чисел (3) на самом деле удовлетворяет системе уравнений (1), т. е. что система (1) совместна. При этом мы используем следующую общеупотребительную символику.

Всякая сумма вида $a_1+a_2+\ldots+a_n$ будет сокращенно обозначаться через $\sum_{i=1}^{n} a_i$. Если же рассматривается сумма, слагаемые которой a_{ij} снабжены двумя индексами, причем $i=1, 2, \ldots, n$, $j=1, 2, \ldots, m$, то можно сначала взять суммы элементов с фиксированным первым индексом, т. е. суммы $\sum_{j=1}^{m} a_{ij}$, где $i=1, 2, \ldots, n$, а затем сложить все эти суммы. Мы получим тогда для суммы всех элементов a_{ij} запись

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}.$$

Можно было бы, однако, вначале складывать слагаемые a_{ij} с фиксированным вторым индексом, а затем уже складывать полученные суммы. Поэтому

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}=\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij},$$

т. е. *в двойной сумме можно менять порядок суммирований.*

Подставим теперь в i-е уравнение системы (1) значения неизвестных (3). Так как левую часть i-го уравнения можно записать в виде $\sum_{j=1}^{n} a_{ij}x_j$ и так как $d_j=\sum_{k=1}^{n} b_kA_{kj}$, то мы получим:

$$\sum_{j=1}^{n} a_{ij}\cdot\frac{d_j}{d}=\frac{1}{d}\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_kA_{kj}\right)=\frac{1}{d}\sum_{k=1}^{n} b_k\left(\sum_{j=1}^{n} a_{ij}A_{kj}\right).$$

Относительно этих преобразований заметим, что число $\frac{1}{d}$ оказалось общим множителем во всех слагаемых и поэтому мы его вынесли за знак суммы; кроме того, после перемены порядка суммирований множитель b_k вынесен за знак внутренней суммы, так как от индекса внутреннего суммирования j он не зависит.

Мы знаем, что выражение $\sum_{j=1}^{n} a_{ij}A_{kj}=a_{i1}A_{k1}+a_{i2}A_{k2}+\ldots+a_{in}A_{kn}$ будет равно d при $k=i$ и равно 0 при всех других k. Таким образом, в нашей внешней сумме по k останется лишь одно слагаемое, а именно b_id, т. е.

$$\sum_{j=1}^{n} a_{ij}\cdot\frac{d_j}{d}=\frac{1}{d}\cdot b_id=b_i.$$

Этим доказано, что система чисел (3) действительно служит решением для системы уравнений (1).

Мы получили следующий важный результат:

Система n линейных уравнений с n неизвестными, определитель которой отличен от нуля, обладает решением, и притом только одним. Это решение получается по формулам (3), т. е. по *правилу Крамера*; формулировка этого правила такова же, как и в случае системы двух уравнений (см. § 2).

Пример. Решить систему линейных уравнений

$$\left.\begin{aligned} 2x_1 + x_2 - 5x_3 + x_4 &= 8, \\ x_1 - 3x_2 \qquad - 6x_4 &= 9, \\ 2x_2 - x_3 + 2x_4 &= -5, \\ x_1 + 4x_2 - 7x_3 + 6x_4 &= 0. \end{aligned}\right\}$$

Определитель этой системы отличен от нуля:

$$d = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27,$$

поэтому к системе применимо правило Крамера. Значения неизвестных будут иметь числителями определители

$$d_1 = \begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix} = 81, \qquad d_2 = \begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix} = -108,$$

$$d_3 = \begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix} = -27, \qquad d_4 = \begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix} = 27.$$

Таким образом,

$$x_1 = 3, \quad x_2 = -4, \quad x_3 = -1, \quad x_4 = 1$$

будет решением нашей системы, притом единственным.

Мы исключили из рассмотрения случай, когда определитель системы n линейных уравнений с n неизвестными (1) равен нулю. Мы отнесем этот случай к гл. 2, где он найдет себе место в общей теории систем с любым числом уравнений и любым числом неизвестных.

Относительно систем n линейных уравнений с n неизвестными сделаем еще одно замечание. Пусть дана система n линейных однородных уравнений с n неизвестными (см. § 1):

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{aligned}\right\} \tag{4}$$

Все определители d_j, $j=1, 2 \ldots, n$, содержат в этом случае столбец, составленный из нулей, и поэтому равны нулю. Таким образом, если определитель системы (4) отличен от нуля, т. е. если к этой системе применимо правило Крамера, то единственным решением системы (4) будет нулевое решение

$$x_1=0, \quad x_2=0, \ \ldots, \ x_n=0. \tag{5}$$

Отсюда вытекает такой результат:

Если система n линейных однородных уравнений с n неизвестными обладает решениями, отличными от нулевого, то определитель этой системы необходимо равен нулю.

В § 12 будет показано, что и обратно, если определитель такой системы действительно равен нулю, то, помимо нулевого решения, существование которого для всякой системы однородных уравнений очевидно, у нее будут существовать и другие решения.

Пример. При каких значениях k система уравнений

$$\left.\begin{aligned} kx_1+x_2=0,\\ x_1+kx_2=0 \end{aligned}\right\}$$

может обладать ненулевыми решениями?

Определитель этой системы

$$\begin{vmatrix} k & 1\\ 1 & k \end{vmatrix}=k^2-1$$

будет равен нулю лишь при $k=\pm 1$. Легко видеть, что при каждом из этих двух значений k заданная система действительно обладает решениями, отличными от нулевого.

Значение правила Крамера заключается главным образом в том, что в тех случаях, когда это правило применимо, оно дает явное выражение для решения системы через коэффициенты этой системы. Практическое использование правила Крамера связано, однако, с весьма громоздкими вычислениями: в случае системы n линейных уравнений с n неизвестными приходится вычислять $n+1$ определитель n-го порядка. Метод последовательного исключения неизвестных, изложенный нами в § 1, является в этом отношении много более удобным, так как вычисления, которых этот метод требует,

по существу равносильны тем, которые приходится выполнять при вычислении одного определителя n-го порядка.

В различных приложениях встречаются системы линейных уравнений, коэффициенты и свободные члены которых являются действительными числами, полученными в результате измерения некоторых физических величин, т. е. известными лишь приближенно, с некоторой точностью. Для решения таких систем изложенные выше методы оказываются иногда неудобными, так как они приводят к результату с плохой точностью, и вместо них разработаны различные итерационные методы, т. е. методы, позволяющие решать указанные системы уравнений при помощи последовательного приближения неизвестных. Изложение этих методов читатель найдет в книгах по теории приближенных вычислений.

ГЛАВА ВТОРАЯ

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ (ОБЩАЯ ТЕОРИЯ)

§ 8. n-мерное векторное пространство

Для построения общей теории систем линейных уравнений недостаточно того аппарата, который с таким успехом послужил нам при решении систем, допускающих применение правила Крамера. Помимо определителей и матриц, мы должны будем использовать одно новое понятие, представляющее, быть может, еще больший общематематический интерес, а именно понятие *многомерного векторного пространства*.

Сначала несколько предварительных замечаний. Из курса аналитической геометрии известно, что всякая точка плоскости определяется (при заданных осях координат) своими двумя координатами, т. е. упорядоченной системой двух действительных чисел; всякий вектор на плоскости определяется своими двумя компонентами, т. е. снова упорядоченной системой двух действительных чисел. Аналогично всякая точка трехмерного пространства определяется своими тремя координатами, всякий вектор в пространстве — тремя компонентами.

В геометрии, а также в механике и физике часто приходится, однако, изучать такие объекты, для задания которых недостаточно трех действительных чисел. Так, рассмотрим совокупность шаров в трехмерном пространстве. Для того чтобы шар был полностью определен, нужно задать координаты его центра и радиус, т. е. задать упорядоченную систему четырех действительных чисел, из которых, впрочем, последнее (радиус) может принимать лишь положительные значения. Рассмотрим, с другой стороны, различные положения твердого тела в пространстве. Положение тела будет вполне определено, если будут указаны координаты его центра тяжести (т. е. три действительных числа), направление некоторой фиксированной оси, проходящей через центр тяжести (два числа — два из трех направляющих косинусов), и, наконец, угол поворота вокруг этой оси. Таким образом, положение твердого тела в пространстве определяется упорядоченной системой из шести действительных чисел.

Эти примеры указывают на целесообразность рассмотрения совокупности всевозможных упорядоченных систем из n действительных чисел. Эта совокупность после введения в нее операций сложения и умножения на число (что будет сделано ниже по аналогии с соответствующими операциями над векторами трехмерного пространства, выраженными через компоненты) и носит название n-мерного векторного пространства. Таким образом, n-мерное пространство есть лишь алгебраическое образование, сохраняющее некоторые простейшие свойства совокупности векторов трехмерного пространства, выходящих из начала координат.

Упорядоченная система n чисел

$$\alpha = (a_1,\ a_2,\ \ldots,\ a_n) \tag{1}$$

называется ***n*-мерным вектором**. Числа a_i, $i = 1, 2, \ldots, n$, будут называться ***компонентами*** вектора α. Векторы α и

$$\beta = (b_1,\ b_2,\ \ldots,\ b_n) \tag{2}$$

будут считаться ***равными*** в том случае, если совпадают их компоненты, стоящие на одинаковых местах, т. е. если $a_i = b_i$ при $i = 1, 2, \ldots, n$. Для обозначения векторов будут употребляться дальше малые греческие буквы, в то время как малые латинские буквы будут использованы для обозначения чисел.

В качестве примеров векторов укажем следующие: 1) Векторы-отрезки, выходящие из начала координат на плоскости или в трехмерном пространстве, будут при фиксированной системе координат соответственно двух- и трехмерными векторами в смысле данного выше определения. 2) Коэффициенты всякого линейного уравнения с n неизвестными составляют n-мерный вектор. 3) Всякое решение любой системы линейных уравнений с n неизвестными будет n-мерным вектором. 4) Если дана матрица из s строк и n столбцов, то ее строки будут n-мерными векторами, столбцы — s-мерными векторами. 5) Сама матрица из s строк и n столбцов может рассматриваться как sn-мерный вектор: достаточно прочесть элементы матрицы подряд, строчку за строчкой; в частности, всякая квадратная матрица порядка n может рассматриваться как n^2-мерный вектор, причем, очевидно, всякий n^2-мерный вектор может быть получен этим путем из некоторой матрицы порядка n.

Суммой векторов (1) и (2) называется вектор

$$\alpha + \beta = (a_1 + b_1,\ a_2 + b_2,\ \ldots,\ a_n + b_n), \tag{3}$$

компоненты которого суть суммы соответствующих компонент слагаемых векторов. Сложение векторов коммутативно и ассоциативно ввиду коммутативности и ассоциативности сложения чисел.

Роль нуля играет ***нулевой вектор***

$$0 = (0,\ 0,\ \ldots,\ 0). \tag{4}$$

Действительно,

$$\alpha + 0 = (a_1 + 0,\ a_2 + 0,\ \ldots,\ a_n + 0) = (a_1,\ a_2,\ \ldots,\ a_n) = \alpha.$$

Для записи нулевого вектора мы употребляем тот же символ 0, как и для числа нуль; решение вопроса, говорится ли в данный момент о числе нуль или о нулевом векторе, никогда не представит затруднений; читатель должен помнить, однако, при изучении ближайших параграфов о возможности различных толкований символа 0.

Назовем *противоположным* вектору (1) вектор

$$-\alpha = (-a_1,\ -a_2,\ \ldots,\ -a_n). \qquad (5)$$

Очевидно, что $\alpha + (-\alpha) = 0$. Теперь легко видеть, что для сложения векторов существует обратная операция — вычитание: *разностью* векторов (1) и (2) будет вектор $\alpha - \beta = \alpha + (-\beta)$, т. е.

$$\alpha - \beta = (a_1 - b_1,\ a_2 - b_2,\ \ldots,\ a_n - b_n). \qquad (6)$$

Сложение n-мерных векторов, определяемое формулой (3), возникло из геометрического сложения векторов на плоскости или в трехмерном пространстве, производимого по правилу параллелограмма. В геометрии используется также умножение вектора на действительное число (на «скаляр»): умножение вектора α на число k означает при $k > 0$ растяжение α в k раз (т. е. сжатие при $k < 1$), а при $k < 0$ растяжение в $|k|$ раз и изменение направления на противоположное. Выражая это правило через компоненты вектора α и переходя к рассматриваемому нами общему случаю, мы получаем такое определение:

Произведением вектора (1) *на число* k называется вектор

$$k\alpha = \alpha k = (ka_1,\ ka_2,\ \ldots,\ ka_n), \qquad (7)$$

компоненты которого равны произведению на k соответственных компонент вектора α.

Из этого определения вытекают следующие важные свойства, проверка которых предоставляется читателю:

$$k(\alpha \pm \beta) = k\alpha \pm k\beta; \qquad (8)$$

$$(k \pm l)\alpha = k\alpha \pm l\alpha; \qquad (9)$$

$$k(l\alpha) = (kl)\alpha; \qquad (10)$$

$$1 \cdot \alpha = \alpha. \qquad (11)$$

Столь же легко проверяются, но могут быть получены и как следствия из свойств (8)—(11), следующие свойства:

$$0 \cdot \alpha = 0; \qquad (12)$$

$$(-1) \cdot \alpha = -\alpha; \qquad (13)$$

$$k \cdot 0 = 0; \qquad (14)$$

$$\text{если } k\alpha = 0, \text{ то или } k = 0, \text{ или } \alpha = 0. \qquad (15)$$

Совокупность всех n-мерных векторов с действительными компонентами, рассматриваемая с определенными в ней операциями сложения векторов и умножения вектора на число, называется *n-мерным векторным пространством.*

Подчеркнем, что в определение n-мерного векторного пространства не входит никакое умножение вектора на вектор. Определить умножение векторов было бы легко — положить, например, что компоненты произведения векторов равны произведениям соответственных компонент сомножителей. Такое умножение не нашло бы у нас, однако, никаких серьезных приложений. Так, векторы-отрезки, выходящие из начала координат на плоскости или в трехмерном пространстве, составляют при фиксированной системе координат двумерное и, соответственно, трехмерное векторные пространства. Сложение векторов и умножение вектора на число имеют в этом примере, как уже отмечено выше, важный геометрический смысл, в то время как покомпонентному умножению векторов нельзя дать никакого разумного геометрического истолкования.

Рассмотрим еще один пример. Левая часть линейного уравнения от n неизвестных, т. е. выражение вида

$$f = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

называется *линейной формой* от неизвестных $x_1, x_2, \ldots, x_n$. Линейная форма f вполне определяется, очевидно, вектором $(a_1, a_2, \ldots, a_n)$ из своих коэффициентов; обратно, всякий n-мерный вектор однозначно определяет некоторую линейную форму. Сложение векторов и умножение вектора на число превращаются в соответствующие операции над линейными формами; эти операции широко использовались нами в § 1. Покомпонентное умножение векторов и в этом примере не имеет никакого смысла.

§ 9. Линейная зависимость векторов

Вектор β из n-мерного векторного пространства называется *пропорциональным* вектору α, если существует такое число k, что $\beta = k\alpha$ (см. формулу (7) предыдущего параграфа). В частности, нулевой вектор пропорционален любому вектору α ввиду равенства $0 = 0 \cdot \alpha$. Если же $\beta = k\alpha$ и $\beta \neq 0$, откуда $k \neq 0$, то $\alpha = k^{-1}\beta$, т. е. для ненулевых векторов пропорциональность обладает свойством симметричности.

Обобщением понятия пропорциональности векторов служит следующее понятие, с которым (для случая строк матрицы) мы уже встречались в § 4: вектор β называется *линейной комбинацией* векторов $\alpha_1, \alpha_2, \ldots, \alpha_s$, если существуют такие числа $l_1, l_2, \ldots, l_s$, что

$$\beta = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_s\alpha_s.$$

Таким образом, j-я компонента вектора β, $j = 1, 2, \ldots, n$, равна,

ввиду определения суммы векторов и произведения вектора на число, сумме произведений j-х компонент векторов $\alpha_1, \alpha_2, \ldots, \alpha_s$ соответственно на $l_1, l_2, \ldots, l_s$.

Система векторов

$$\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_r \qquad (r \geqslant 2) \tag{1}$$

называется *линейно зависимой*, если хотя бы один из этих векторов является линейной комбинацией остальных векторов системы (1), и *линейно независимой* — в противоположном случае.

Укажем другую форму этого весьма важного определения: система векторов (1) линейно зависима, если существуют такие числа $k_1, k_2, \ldots, k_r$, *хотя бы одно из которых отлично от нуля*, что имеет место равенство

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0. \tag{2}$$

Доказательство эквивалентности этих двух определений не представляет затруднений. Пусть, например, вектор α_r из системы (1) есть линейная комбинация остальных векторов:

$$\alpha_r = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1}.$$

Отсюда вытекает равенство

$$l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1} - \alpha_r = 0,$$

т. е. равенство вида (2), где $k_i = l_i$ для $i = 1, 2, \ldots, r-1$ и $k_r = -1$, т. е. $k_r \neq 0$. Пусть, обратно, векторы (1) связаны соотношением (2), в котором, например, $k_r \neq 0$. Тогда

$$\alpha_r = \left(-\frac{k_1}{k_r}\right)\alpha_1 + \left(-\frac{k_2}{k_r}\right)\alpha_2 + \ldots + \left(-\frac{k_{r-1}}{k_r}\right)\alpha_{r-1},$$

т. е. вектор α_r оказался линейной комбинацией векторов $\alpha_1, \alpha_2, \ldots, \alpha_{r-1}$.

Пример. Система векторов

$$\alpha_1 = (5, 2, 1), \quad \alpha_2 = (-1, 3, 3), \quad \alpha_3 = (9, 7, 5), \quad \alpha_4 = (3, 8, 7)$$

линейно зависима, так как векторы связаны соотношением

$$4\alpha_1 - \alpha_2 - 3\alpha_3 + 2\alpha_4 = 0.$$

В этом соотношении все коэффициенты отличны от нуля. Между нашими векторами существуют, однако, и другие линейные зависимости, в которых некоторые из коэффициентов равны нулю, например

$$2\alpha_1 + \alpha_2 - \alpha_3 = 0, \quad 3\alpha_2 + \alpha_3 - 2\alpha_4 = 0.$$

Второе из данных выше определений линейной зависимости применимо и к случаю $r = 1$, т. е. к случаю системы, состоящей из одного вектора α: *эта система тогда и только тогда будет линейно зависимой, если* $\alpha = 0$. Действительно, если $\alpha = 0$, то, например, при $k = 1$ будет $k\alpha = 0$. Обратно, если $k\alpha = 0$ и $k \neq 0$, то $\alpha = 0$.

Отметим следующее свойство понятия линейной зависимости. *Если некоторая подсистема системы векторов* (1) *линейно зависима, то и вся система* (1) *линейно зависима.*

Действительно, пусть векторы $\alpha_1, \alpha_2, \ldots, \alpha_s$ из системы (1), где $s<r$, связаны соотношением

$$k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s=0,$$

в котором не все коэффициенты равны нулю. Отсюда следует соотношение

$$k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s+0\cdot\alpha_{s+1}+\ldots+0\cdot\alpha_r=0,$$

т. е. система (1) линейно зависима.

Из этого свойства вытекает линейная зависимость всякой системы векторов, содержащей два равных или, вообще, два пропорциональных вектора, а также всякой системы, содержащей нулевой вектор. Заметим, что доказанному сейчас свойству можно дать такую формулировку: *если система векторов* (1) *линейно независима, то и всякая ее подсистема также линейно независима.*

Возникает вопрос, как много векторов может содержать линейно независимая система n-мерных векторов и существуют ли, в частности, такие системы с произвольно большим числом векторов. Для ответа на этот вопрос рассмотрим в n-мерном векторном пространстве векторы

$$\left.\begin{aligned} \varepsilon_1&=(1,\ 0,\ 0,\ \ldots,\ 0),\\ \varepsilon_2&=(0,\ 1,\ 0,\ \ldots,\ 0),\\ &\cdots\cdots\cdots\cdots\\ \varepsilon_n&=(0,\ 0,\ 0,\ \ldots,\ 1), \end{aligned}\right\} \qquad (3)$$

называемые *единичными векторами* этого пространства. *Система единичных векторов будет линейно независимой*: пусть

$$k_1\varepsilon_1+k_2\varepsilon_2+\ldots+k_n\varepsilon_n=0;$$

так как левая часть этого равенства равна вектору $(k_1, k_2, \ldots, k_n)$, то

$$(k_1, k_2, \ldots, k_n)=0,$$

т. е. $k_i=0$, $i=1, 2, \ldots, n$, так как все компоненты нулевого вектора равны нулю, а равенство векторов равносильно равенству их соответственных компонент.

Мы нашли, таким образом, в n-мерном векторном пространстве одну линейно независимую систему, состоящую из n векторов. Читатель увидит позже, что на самом деле в этом пространстве существует бесконечно много различных таких систем.

Докажем, с другой стороны, следующую теорему:

Всякие s векторов n-мерного векторного пространства составляют при $s>n$ линейно зависимую систему.

Пусть, в самом деле, нам даны векторы

$$\begin{aligned}\alpha_1&=(a_{11}, a_{12}, \ldots, a_{1n}),\\ \alpha_2&=(a_{21}, a_{22}, \ldots, a_{2n}),\\ &\cdots\cdots\cdots\\ \alpha_s&=(a_{s1}, a_{s2}, \ldots, a_{sn}).\end{aligned}$$

Нам нужно подобрать такие числа $k_1, k_2, \ldots, k_s$, не все равные нулю, что

$$k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s=0. \tag{4}$$

Переходя от равенства (4) к соответствующим равенствам между компонентами, получаем

$$\left.\begin{aligned}a_{11}k_1+a_{21}k_2+\ldots+a_{s1}k_s&=0,\\ a_{12}k_1+a_{22}k_2+\ldots+a_{s2}k_s&=0,\\ \cdots\cdots\cdots&\cdots\\ a_{1n}k_1+a_{2n}k_2+\ldots+a_{sn}k_s&=0.\end{aligned}\right\} \tag{5}$$

Равенства (5) составляют, однако, систему n линейных однородных уравнений относительно s неизвестных $k_1, k_2, \ldots, k_s$. Число уравнений в этой системе меньше числа неизвестных, а поэтому, как доказано в конце § 1, эта система обладает ненулевыми решениями. Таким образом, можно подобрать числа $k_1, k_2, \ldots, k_s$, не все равные нулю, которые удовлетворяют требованию (4). Теорема доказана.

Назовем линейно независимую систему n-мерных векторов

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{6}$$

максимальной линейно независимой системой, если добавление к этой системе любого n-мерного вектора β дает уже линейно зависимую систему. Так как во всякой линейной зависимости, связывающей векторы $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta$, коэффициент при β должен быть отличным от нуля — иначе система (6) была бы линейно зависимой, — то вектор β линейно выражается через векторы (6). Поэтому система векторов (6) тогда и только тогда будет максимальной линейно независимой системой, если векторы (6) линейно независимы, а любой n-мерный вектор β является их линейной комбинацией.

Из результатов, полученных выше, вытекает, что *в n-мерном пространстве всякая линейно независимая система, состоящая из n векторов, будет максимальной,* а также что *любая максимальная линейно независимая система векторов этого пространства состоит не более чем из n векторов.*

Всякая линейно независимая система n-мерных векторов содержится хотя бы в одной максимальной линейно независимой

мой системе. В самом деле, если заданная система векторов не максимальна, то к ней можно добавить один вектор так, что полученная система останется линейно независимой. Если эта новая система все еще не максимальна, то к ней можно добавить еще один вектор, и т. д. Этот процесс не может, однако, продолжаться бесконечно, так как уже любая система n-мерных векторов, состоящая из $n+1$ вектора, будет линейно зависимой.

Так как всякая система, состоящая из одного ненулевого вектора, линейно независима, то мы получаем, что *всякий ненулевой вектор содержится в некоторой максимальной линейно независимой системе*, а поэтому *в n-мерном векторном пространстве существует бесконечно много различных максимальных линейно независимых систем векторов*.

Возникает вопрос, существуют ли в этом пространстве максимальные линейно независимые системы с меньшим, чем n, числом векторов или же число векторов в любой такой системе непременно равно n? Ответ на этот важный вопрос будет дан ниже, после некоторых предварительных рассмотрений.

Если вектор β является линейной комбинацией векторов

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_r, \tag{7}$$

то часто говорят, что β *линейно выражается через систему* (7). Понятно, что если вектор β линейно выражается через некоторую подсистему этой системы, то он будет линейно выражаться и через систему (7) — достаточно остальные векторы системы взять с коэффициентами, равными нулю. Обобщая эту терминологию, говорят, что *система векторов*

$$\beta_1,\ \beta_2,\ \ldots,\ \beta_s \tag{8}$$

линейно выражается через систему (7), если всякий вектор β_i, $i=1, 2, \ldots, s$, является линейной комбинацией векторов системы (7).

Докажем транзитивность этого понятия: *если система* (8) *линейно выражается через систему* (7), *а система векторов*

$$\gamma_1,\ \gamma_2,\ \ldots,\ \gamma_t \tag{9}$$

линейно выражается через систему (8), *то* (9) *будет линейно выражаться и через* (7).

В самом деле,

$$\gamma_j=\sum_{i=1}^{s} l_{ji}\beta_i,\quad j=1,\ 2,\ \ldots,\ t, \tag{10}$$

но $\beta_i=\sum_{m=1}^{r} k_{im}\alpha_m$, $i=1, 2, \ldots, s$. Подставляя эти выражения в (10), получаем:

$$\gamma_j=\sum_{i=1}^{s} l_{ji}\left(\sum_{m=1}^{r} k_{im}\alpha_m\right)=\sum_{m=1}^{r}\left(\sum_{i=1}^{s} l_{ji}k_{im}\right)\alpha_m,$$

т. е. всякий вектор γ_j, $j=1, 2, \ldots, t$, будет линейной комбинацией векторов системы (7).

Две системы векторов называются *эквивалентными*, если каждая из них линейно выражается через другую. Из доказанной сейчас транзитивности свойства систем векторов линейно выражаться друг через друга вытекает транзитивность понятия эквивалентности систем векторов, а также следующее утверждение: *если две системы векторов эквивалентны и если некоторый вектор линейно выражается через одну из этих систем, то он будет линейно выражаться и через другую.*

Нельзя утверждать, что если одна из двух эквивалентных между собой систем векторов линейно независима, то этим же свойством обладает и другая система. Если же обе эти системы линейно независимы, то о числе векторов, входящих в них, можно сделать одно важное высказывание. Докажем сначала следующую теорему, которую ввиду ее роли в дальнейшем и для удобства ссылок назовем основной теоремой.

Если в n-мерном векторном пространстве даны две системы векторов:

$$\text{(I)} \qquad \alpha_1, \alpha_2, \ldots, \alpha_r,$$

$$\text{(II)} \qquad \beta_1, \beta_2, \ldots, \beta_s,$$

из которых первая линейно независима и линейно выражается через вторую, то число векторов в первой системе не больше, чем во второй, т. е. $r \leqslant s$.

Пусть, в самом деле, $r > s$. По условию, каждый вектор системы (I) линейно выражается через систему (II):

$$\left.\begin{aligned} \alpha_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \ldots + a_{1s}\beta_s, \\ \alpha_2 &= a_{21}\beta_1 + a_{22}\beta_2 + \ldots + a_{2s}\beta_s, \\ &\ldots\ldots\ldots\ldots\ldots \\ \alpha_r &= a_{r1}\beta_1 + a_{r2}\beta_2 + \ldots + a_{rs}\beta_s. \end{aligned}\right\} \qquad (11)$$

Коэффициенты этих линейных выражений составляют систему из r s-мерных векторов:

$$\begin{aligned} \gamma_1 &= (a_{11}, a_{12}, \ldots, a_{1s}), \\ \gamma_2 &= (a_{21}, a_{22}, \ldots, a_{2s}), \\ &\ldots\ldots\ldots\ldots \\ \gamma_r &= (a_{r1}, a_{r2}, \ldots, a_{rs}). \end{aligned}$$

Так как $r > s$, то эти векторы линейно зависимы, т. е.

$$k_1\gamma_1 + k_2\gamma_2 + \ldots + k_r\gamma_r = 0,$$

где не все коэффициенты k_1, k_2 ..., k_r равны нулю. Отсюда мы приходим к следующим равенствам между компонентами:

$$\sum_{i=1}^{r} k_i a_{ij} = 0, \qquad j = 1,\ 2,\ \ldots,\ s. \tag{12}$$

Рассмотрим теперь следующую линейную комбинацию векторов системы (I):

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r$$

или, короче, $\sum_{i=1}^{r} k_i\alpha_i$. Используя (11) и (12), получаем:

$$\sum_{i=1}^{r} k_i\alpha_i = \sum_{i=1}^{r} k_i \left(\sum_{j=1}^{s} a_{ij}\beta_j \right) = \sum_{j=1}^{s} \left(\sum_{i=1}^{r} k_i a_{ij} \right) \beta_j = 0;$$

это противоречит, однако, линейной независимости системы (I).

Из доказанной сейчас основной теоремы вытекает следующий результат:

Всякие две эквивалентные линейно независимые системы векторов содержат равное число векторов.

Любые две максимальные линейно независимые системы n-мерных векторов будут, очевидно, эквивалентными. Они состоят, следовательно, из одного и того же числа векторов, а так как существуют, как нам известно, системы такого рода, состоящие из n векторов, то мы получаем, наконец, ответ на поставленный ранее вопрос: *всякая максимальная линейно независимая система векторов n-мерного векторного пространства состоит из n векторов.*

Из полученных результатов можно вывести и другие следствия.

Если в данной линейно зависимой системе векторов взяты две в ней максимальные линейно независимые подсистемы, т. е. такие подсистемы, к которым нельзя присоединить ни одного вектора нашей системы, не нарушая линейной независимости, то эти подсистемы содержат равное число векторов.

В самом деле, если в системе векторов

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_r \tag{13}$$

подсистема

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_s,\quad s < r, \tag{14}$$

будет максимальной линейно независимой подсистемой, то всякий из векторов $\alpha_{s+1}, \ldots, \alpha_r$ будет линейно выражаться через систему (14). С другой стороны, всякий вектор α_i из системы (13) линейно выражается через эту систему: достаточно взять при самом векторе α_i коэффициент 1, а при всех остальных векторах системы коэффициент 0. Теперь легко видеть, что системы (13) и (14) эквивалентны. Отсюда следует, что система (13) эквивалентна всякой из своих максимальных линейно независимых подсистем, а поэтому все

эти подсистемы эквивалентны между собой, т. е., будучи линейно независимыми, содержат по одному и тому же числу векторов.

Число векторов, входящих в любую максимальную линейно независимую подсистему данной системы векторов, называется *рангом* этой системы. Используя это понятие, выведем еще одно следствие из основной теоремы.

Пусть даны две системы n-мерных векторов:

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{15}$$

и

$$\beta_1, \beta_2, \ldots, \beta_s, \tag{16}$$

не обязательно линейно независимые, причем ранг системы (15) *равен числу k, ранг системы* (16) — *числу l. Если первая система линейно выражается через вторую, то $k \leqslant l$. Если же эти системы эквивалентны, то $k = l$.*

В самом деле, пусть

$$\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k} \tag{17}$$

и

$$\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_l} \tag{18}$$

будут, соответственно, любые максимальные линейно независимые подсистемы систем (15) и (16). Тогда системы (15) и (17) эквивалентны между собой, и это же верно для систем (16) и (18). Из того, что система (15) линейно выражается через систему (16), вытекает теперь, что система (17) также линейно выражается через систему (16), а поэтому и через эквивалентную ей систему (18), после чего остается, используя линейную независимость системы (17), применить основную теорему. Второе утверждение доказываемого следствия непосредственно вытекает из первого.

§ 10. Ранг матрицы

Если дана некоторая система n-мерных векторов, то возникает естественный вопрос, является ли эта система векторов линейно зависимой или нет. Нельзя рассчитывать на то, что в каждом конкретном случае решение этого вопроса будет получено без затруднений: при поверхностном рассмотрении системы векторов

$$\alpha = (2, -5, 1, -1), \quad \beta = (1, 3, 6, 5), \quad \gamma = (-1, 4, 1, 2)$$

трудно заметить в ней какие-либо линейные зависимости, хотя в действительности эти векторы связаны соотношением

$$7\alpha - 3\beta + 11\gamma = 0$$

Один метод для решения этого вопроса дает § 1; так как компоненты заданных векторов нам известны, то, считая неизвестными

коэффициенты искомой линейной зависимости, мы получаем систему линейных однородных уравнений, которую и решаем методом Гаусса. В настоящем параграфе будет указан другой подход к рассматриваемому вопросу; одновременно мы значительно приблизимся к нашей основной цели — решению произвольных систем линейных уравнений.

Пусть дана матрица

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{s1} & a_{s2} & \dots & a_{sn} \end{pmatrix},$$

содержащая s строк и n столбцов, причем числа s и n никак не связаны между собой. Столбцы этой матрицы, рассматриваемые как s-мерные векторы, могут, вообще говоря, быть линейно зависимыми. Ранг системы столбцов, т. е. максимальное число линейно независимых столбцов матрицы A (точнее, число столбцов, входящих в любую максимальную линейно независимую подсистему системы столбцов), называется *рангом* этой матрицы.

Понятно, что подобным же образом строки матрицы A можно рассматривать как n-мерные векторы. Оказывается, что ранг системы строк матрицы равен рангу системы ее столбцов, т. е. равен рангу этой матрицы. Доказательство этого весьма неожиданного утверждения будет получено после того, как мы укажем еще одну форму определения ранга матрицы, дающую заодно способ его практического вычисления.

Обобщим сначала на случай прямоугольных матриц понятие минора. Выбираем в матрице A произвольные k строк и k столбцов, $k \leqslant \min(s, n)$. Элементы, стоящие на пересечении этих строк и столбцов, составляют квадратную матрицу k-го порядка, определитель которой называется *минором k-го порядка* матрицы A. Дальше нас будут интересовать порядки тех миноров матрицы A, которые отличны от нуля, а именно наивысший среди этих порядков. При его разыскании полезно учитывать следующее замечание: *если все миноры k-го порядка матрицы A равны нулю, то равны нулю и все миноры более высоких порядков.* В самом деле, разлагая всякий минор порядка $k+j$, $k<k+j\leqslant \min(s, n)$, на основании теоремы Лапласа, по любым k строкам, мы представим этот минор в виде суммы миноров порядка k, умноженных на некоторые миноры порядка j, и этим докажем, что он равен нулю.

Докажем теперь следующую теорему о ранге матрицы: *Наивысший порядок отличных от нуля миноров матрицы A равен рангу этой матрицы.*

Доказательство. Пусть наивысший порядок отличных от нуля миноров матрицы A равен r. Предположим, — что не нарушает

общности доказательства,— что минор r-го порядка D, стоящий в левом верхнем углу матрицы

$$A=\begin{pmatrix} \begin{array}{|ccc|} a_{11} & \dots & a_{1r} \\ \dots & D & \dots \\ a_{r1} & \dots & a_{rr} \\ \hline \end{array} & \begin{array}{ccc} a_{1,r+1} & \dots & a_{1n} \\ . & . & . \\ a_{r,r+1} & \dots & a_{rn} \end{array} \\ \begin{array}{ccc} a_{r+1,1} & \dots & a_{r+1,r} \\ . & . & . \\ a_{s1} & \dots & a_{sr} \end{array} & \begin{array}{ccc} a_{r+1,r+1} & \dots & a_{r+1,n} \\ . & . & . \\ a_{s,r+1} & \dots & a_{sn} \end{array} \end{pmatrix},$$

отличен от нуля, $D \neq 0$. Тогда первые r столбцов матрицы A будут между собой линейно независимыми: если бы между ними существовала линейная зависимость, то, так как при сложении векторов складываются соответствующие компоненты, между столбцами минора D существовала бы эта же линейная зависимость и поэтому минор D был бы равен нулю.

Докажем теперь, что всякий l-й столбец матрицы A, $r < l \leqslant n$, будет линейной комбинацией первых r столбцов. Берем любое i, $1 \leqslant i \leqslant s$, и строим вспомогательный определитель $(r+1)$-го порядка

$$\Delta_i=\begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1l} \\ . & . & . & . \\ a_{r1} & \dots & a_{rr} & a_{rl} \\ a_{i1} & \dots & a_{ir} & a_{il} \end{vmatrix},$$

получающийся «окаймлением» минора D соответствующими элементами l-го столбца и i-й строки. При любом i определитель Δ_i равен нулю. Действительно, если $i > r$, то Δ_i будет минором $(r+1)$-го порядка нашей матрицы A и поэтому равен нулю ввиду выбора числа r. Если же $i \leqslant r$, то Δ_i уже не будет минором матрицы A, так как не может быть получен вычеркиванием из этой матрицы некоторых ее строк и столбцов; однако определитель Δ_i будет содержать теперь две равные строки и, следовательно, снова равен нулю.

Рассмотрим алгебраические дополнения элементов последней строки определителя Δ_i. Алгебраическим дополнением для элемента a_{il} служит, очевидно, минор D. Если же $1 \leqslant j \leqslant r$, то алгебраическим дополнением для элемента a_{ij} в Δ_i будет число

$$A_j=(-1)^{(r+1)+j}\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1r} & a_{1l} \\ . & . & . & . & . & . & . \\ a_{r1} & \dots & a_{r,j-1} & a_{r,j+1} & \dots & a_{rr} & a_{rl} \end{vmatrix};$$

оно не зависит от i и поэтому обозначено через A_j. Таким образом, разлагая определитель Δ_i по его последней строке и приравнивая это разложение нулю, так как $\Delta_i=0$, мы получим:

$$a_{i1}A_1+a_{i2}A_2+\dots+a_{ir}A_r+a_{il}D=0,$$

откуда, ввиду $D \neq 0$,

$$a_{il} = -\frac{A_1}{D} a_{i1} - \frac{A_2}{D} a_{i2} - \ldots - \frac{A_r}{D} a_{ir}.$$

Это равенство справедливо при всех i, $i = 1, 2, \ldots, s$, а так как его коэффициенты от i не зависят, то мы получаем, что весь l-й столбец матрицы A будет суммой ее первых r столбцов, взятых, соответственно, с коэффициентами $-\frac{A_1}{D}, -\frac{A_2}{D}, \ldots, -\frac{A_r}{D}$.

Таким образом, в системе столбцов матрицы A мы нашли максимальную линейно независимую подсистему, состоящую из r столбцов. Этим доказано, что ранг матрицы A равен r, т. е. доказана теорема о ранге.

Эта теорема дает метод для практического вычисления ранга матрицы, а поэтому и для решения вопроса о существовании линейной зависимости в данной системе векторов; составляя матрицу, для которой данные векторы служат столбцами, и вычисляя ранг этой матрицы, мы находим максимальное число линейно независимых векторов нашей системы.

Метод нахождения ранга матрицы, основанный на теореме о ранге, требует вычисления хотя и конечного, но, быть может, очень большого числа миноров этой матрицы. Следующее замечание позволяет, однако, внести в этот метод значительные упрощения. Если читатель просмотрит еще раз доказательство теоремы о ранге, то заметит, что мы не использовали при его проведении равенства нулю всех миноров $(r+1)$-го порядка матрицы A — в действительности употреблялись лишь те миноры $(r+1)$-го порядка, которые окаймляют данный не равный нулю минор r-го порядка D (т. е. содержат его целиком внутри себя), и поэтому из равенства нулю лишь этих миноров вытекает, что r есть максимальное число линейно независимых столбцов матрицы A; последнее же влечет за собой равенство нулю вообще всех миноров $(r+1)$-го порядка этой матрицы. Мы приходим к следующему правилу вычисления ранга матрицы:

При вычислении ранга матрицы следует переходить от миноров меньших порядков к минорам бо́льших порядков. Если уже найден минор k-го порядка D, отличный от нуля, то требуют вычисления лишь миноры $(k+1)$-го порядка, окаймляющие минор D: если все они равны нулю, то ранг матрицы равен k.

Примеры.

1. Найти ранг матрицы

$$A = \begin{pmatrix} 2 & -4 & 3 & 1 & 0 \\ 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 4 & -7 & 4 & -4 & 5 \end{pmatrix}.$$

Минор второго порядка, стоящий в левом верхнем углу этой матрицы, равен нулю. Однако в матрице содержатся и отличные от нуля миноры второго порядка, например

$$d=\begin{vmatrix} -4 & 3 \\ -2 & 1 \end{vmatrix} \neq 0.$$

Минор третьего порядка

$$d'=\begin{vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{vmatrix},$$

окаймляющий минор d, отличен от нуля, $d'=1$, однако оба минора четвертого порядка, окаймляющие минор d', равны нулю:

$$\begin{vmatrix} 2 & -4 & 3 & 1 \\ 1 & -2 & 1 & -4 \\ 0 & 1 & -1 & 3 \\ 4 & -7 & 4 & -4 \end{vmatrix}=0, \qquad \begin{vmatrix} 2 & -4 & 3 & 0 \\ 1 & -2 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 4 & -7 & 4 & 5 \end{vmatrix}=0.$$

Таким образом, ранг матрицы A равен трем.

2. Найти максимальную линейно независимую подсистему в системе векторов

$$\alpha_1=(2,\ -2,\ -4),\quad \alpha_2=(1,\ 9,\ 3),\quad \alpha_3=(-2,\ -4,\ 1),\quad \alpha_4=(3,\ 7,\ -1).$$

Составляем матрицу

$$\begin{pmatrix} 2 & 1 & -2 & 3 \\ -2 & 9 & -4 & 7 \\ -4 & 3 & 1 & -1 \end{pmatrix},$$

для которой данные векторы служат столбцами. Ранг этой матрицы равен двум: минор второго порядка, стоящий в левом верхнем углу, отличен от нуля, но оба минора третьего порядка, его окаймляющие, равны нулю. Отсюда следует, что векторы α_1, α_2 составляют в заданной системе одну из максимальных линейно независимых подсистем.

В качестве следствия из теоремы о ранге матрицы докажем утверждение, уже высказанное ранее:

Максимальное число линейно независимых строк всякой матрицы равно максимальному числу ее линейно независимых столбцов, т. е. равно рангу этой матрицы.

Для доказательства транспонируем матрицу, т. е. сделаем ее строки столбцами, сохраняя их нумерацию. При транспонировании максимальный порядок отличных от нуля миноров матрицы не может измениться, так как транспонирование не меняет определителя, а для всякого минора исходной матрицы минор, полученный из него транспонированием, содержится в новой матрице, и обратно. Отсюда следует, что ранг новой матрицы равен рангу исходной матрицы; он равен, вместе с тем, максимальному числу линейно независимых столбцов новой матрицы, т. е. максимальному числу линейно независимых с т р о к исходной матрицы.

Пример. В § 8 уже было введено понятие линейной формы от n неизвестных и определено сложение линейных форм и их умножение на число. Это определение позволяет перенести на линейные формы понятие линейной зависимости со всеми его свойствами.

Пусть дана система линейных форм

$$\begin{aligned} f_1 &= x_1 + 2x_2 + x_3 + 3x_4, \\ f_2 &= 4x_1 - x_2 - 5x_3 - 6x_4, \\ f_3 &= x_1 - 3x_2 - 4x_3 - 7x_4, \\ f_4 &= 2x_1 + x_2 - x_3. \end{aligned}$$

Нужно выделить в ней максимальную линейно независимую подсистему.

Составим матрицу из коэффициентов этих форм:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 4 & -1 & -5 & -6 \\ 1 & -3 & -4 & -7 \\ 2 & 1 & -1 & 0 \end{pmatrix}$$

и найдем ее ранг. Минор второго порядка, стоящий в левом верхнем углу, отличен от нуля, но, как легко проверить, все четыре минора третьего порядка, его окаймляющие, равны нулю. Отсюда следует, что первые две строки нашей матрицы линейно независимы, а третья и четвертая будут их линейными комбинациями. Система f_1, f_2 будет, следовательно, искомой подсистемой заданной системы линейных форм.

Укажем еще одно важное следствие из теоремы о ранге матрицы.

Определитель n-го порядка тогда и только тогда равен нулю, если между его строками существует линейная зависимость.

В одну сторону это утверждение уже доказано в § 4 (свойство 8). Пусть теперь нам дан определитель n-го порядка, равный нулю, т. е. дана, иными словами, квадратная матрица n-го порядка, единственный минор которой, имеющий максимальный порядок, равен нулю. Отсюда следует, что наивысший порядок отличных от нуля миноров этой матрицы меньше n, т. е. ранг меньше n, а поэтому, на основании доказанного выше, строки этой матрицы линейно зависимы.

Понятно, что в формулировке доказанного сейчас следствия можно вместо строк говорить о столбцах определителя.

Для вычисления ранга матрицы существует еще один метод, не связанный с теоремой о ранге и не требующий вычисления определителей. Он применим, впрочем, только в том случае, если мы хотим знать лишь самый ранг и не интересуемся тем, какие именно столбцы (или строки) составляют максимальную линейно независимую систему. Изложим этот метод.

Элементарными преобразованиями матрицы A называются следующие преобразования этой матрицы:

(a) перемена мест (транспозиция) двух строк или двух столбцов;

(b) умножение строки (или столбца) на произвольное отличное от нуля число;

(c) прибавление к одной строке (или столбцу) другой строки (столбца), умноженной на некоторое число.

Легко видеть, что *элементарные преобразования не меняют ранга матрицы*. Действительно, если эти преобразования применяются, например, к столбцам матрицы, то система столбцов, рассматриваемых как векторы, заменяется эквивалентной системой. Докажем это лишь для преобразования (c), так как для (a) и (b) это очевидно. Пусть к i-му столбцу прибавляется j-й столбец, умноженный на число k. Если столбцами матрицы до преобразования служили векторы

$$\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_n, \tag{1}$$

то после преобразования столбцами матрицы будут векторы

$$\alpha_1, \ldots, \alpha_i' = \alpha_i + k\alpha_j, \ldots, \alpha_j, \ldots, \alpha_n. \tag{2}$$

Система (2) линейно выражается через систему (1), а равенство

$$\alpha_i = \alpha_i' - k\alpha_j$$

показывает, что система (1), в свою очередь, линейно выражается через (2). Эти системы, следовательно, эквивалентны, и поэтому их максимальные линейно независимые подсистемы состоят из одинакового числа векторов.

Таким образом, при вычислении ранга матрицы можно предварительно ее упростить при помощи некоторой комбинации элементарных преобразований.

Говорят, что матрица, содержащая s строк и n столбцов, имеет *диагональную форму*, если все ее элементы равны нулю, кроме элементов a_{11}, $a_{22}, \ldots, a_{rr}$ (где $0 \leqslant r \leqslant \min(s, n)$), равных единице. Ранг этой матрицы равен, очевидно, r.

Всякую матрицу можно элементарными преобразованиями привести к диагональной форме.

В самом деле, пусть дана матрица

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ . & . & . \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}.$$

Если все ее элементы равны нулю, то она уже имеет диагональную форму. Если же в ней есть элементы, отличные от нуля, то транспозицией строк и столбцов можно добиться того, чтобы элемент a_{11} был отличен от нуля. Умножая затем первую строку на a_{11}^{-1}, мы превратим элемент a_{11} в единицу. Если мы вычтем теперь из j-го столбца, $j > 1$, первый столбец, умноженный на a_{1j}, то элемент a_{1j} будет заменен нулем. Делая это преобразование со всеми столбцами, начиная со второго, а также со всеми строками, мы придем к матрице вида

$$A' = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & a_{22}' & \ldots & a_{2n}' \\ . & . & . & . \\ 0 & a_{s2}' & \ldots & a_{sn}' \end{pmatrix}.$$

Совершая такие же преобразования с матрицей, остающейся в правом нижнем углу, и т. д., мы после конечного числа шагов придем к диагональной матрице, имеющей тот же ранг, что и исходная матрица A.

Таким образом, *для нахождения ранга матрицы нужно элементарными преобразованиями привести эту матрицу к диагональной форме и подсчитать число единиц, стоящих в последней на главной диагонали.*

Пример. Найти ранг матрицы

$$A=\begin{pmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \\ 2 & 3 & 0 \end{pmatrix}.$$

Переставляя в этой матрице первый и второй столбец, а затем умножая первую строку на число $\frac{1}{2}$, мы придем к матрице

$$\begin{pmatrix} 1 & 0 & -2 \\ -4 & -1 & 5 \\ 1 & 3 & 7 \\ 5 & 0 & -10 \\ 3 & 2 & 0 \end{pmatrix}.$$

Прибавляя к ее третьему столбцу удвоенный первый столбец, а затем прибавляя некоторое кратное новой первой строки к каждой из остальных строк, мы получим матрицу

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \\ 0 & 2 & 6 \end{pmatrix}.$$

Умножая, наконец, вторую строку на -1, вычитая из третьего столбца утроенный второй столбец, а затем вычитая из третьей и пятой строк некоторые кратные новой второй строки, мы придем к искомой диагональной форме

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Таким образом, ранг матрицы A равен двум.

В гл. 13 мы еще раз встретимся с элементарными преобразованиями и диагональной формой матриц; это будут, впрочем, матрицы, элементами которых являются не числа, а многочлены.

§ 11. Системы линейных уравнений

Мы переходим к изучению произвольных систем линейных уравнений, причем уже не делаем предположения, что число уравнений системы равно числу неизвестных. Наши результаты будут, впрочем, применимы и к тому случаю (оставленному в § 7 без рассмотрения), когда число уравнений равно числу неизвестных, но определитель системы равен нулю.

Пусть дана система линейных уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots& \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n &= b_s. \end{aligned}\right\} \qquad (1)$$

Как мы знаем из § 1, прежде всего следует решить вопрос о совместности этой системы. Для этой цели возьмем матрицу A из коэффициентов системы и «расширенную» матрицу $\bar{A}$, полученную присоединением к A столбца из свободных членов,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} & b_s \end{pmatrix},$$

и вычислим ранги этих матриц. Легко видеть, что *ранг матрицы $\bar{A}$ либо равен рангу матрицы A, либо на единицу больше последнего.* В самом деле, берем некоторую максимальную линейно независимую систему столбцов матрицы A. Она будет линейно независимой и в матрице $\bar{A}$. Если она сохраняет и свойство максимальности, т. е. столбец из свободных членов через нее линейно выражается, то ранги матриц A и $\bar{A}$ равны; в противоположном случае, присоединяя к этой системе столбец из свободных членов, мы получаем линейно независимую систему столбцов матрицы $\bar{A}$, которая будет в ней максимальной.

Вопрос о совместности системы линейных уравнений полностью решается следующей теоремой.

Теорема Кронекера—Капелли. *Система линейных уравнений* (1) *тогда и только тогда совместна, когда ранг расширенной матрицы $\bar{A}$ равен рангу матрицы A.*

Доказательство. 1. Пусть система (1) совместна и пусть $k_1, k_2, \ldots, k_n$ будет одним из ее решений. Подставляя эти числа вместо неизвестных в систему (1), мы получим s тождеств, которые показывают, что последний столбец матрицы $\bar{A}$ является суммой всех остальных столбцов, взятых соответственно с коэффициентами $k_1, k_2, \ldots, k_n$. Всякий другой столбец матрицы $\bar{A}$ входит и в матрицу A и поэтому линейно выражается через все столбцы этой матрицы. Обратно, всякий столбец матрицы A является столбцом и в $\bar{A}$, т. е. линейно выражается через столбцы этой матрицы. Отсюда следует, что системы столбцов матриц A и $\bar{A}$ эквивалентны между собой, а поэтому, как доказано в конце § 9, обе эти системы s-мерных векторов имеют один и тот же ранг; иными словами, ранги матриц A и $\bar{A}$ равны между собой.

2. Пусть теперь дано, что матрицы A и $\bar{A}$ имеют равные ранги. Отсюда следует, что любая максимальная линейно независимая система столбцов матрицы A остается максимальной линейно независимой системой и в матрице $\bar{A}$. Таким образом, через эту систему, а поэтому и вообще через систему столбцов матрицы A, линейно выражается последний столбец матрицы $\bar{A}$. Существует, следовательно, такая система коэффициентов $k_1, k_2, \ldots, k_n$, что сумма столбцов матрицы A, взятых с этими коэффициентами, равна столбцу из свободных членов, а потому числа $k_1, k_2, \ldots, k_n$ составляют решение системы (1). Таким образом, совпадение рангов матриц A и $\bar{A}$ влечет за собой совместность системы (1).

Теорема полностью доказана. При ее применении к конкретным примерам необходимо вычислить сперва ранг матрицы A, для чего найти один из тех отличных от нуля миноров этой матрицы, что все миноры, его окаймляющие, равны нулю; пусть это будет минор M. После этого следует вычислить все миноры матрицы $\bar{A}$, окаймляющие M, но в A не содержащиеся (так называемые *характеристические определители* системы (1)). Если они все равны нулю, то ранг матрицы $\bar{A}$ равен рангу матрицы A и потому система (1) совместна, в противоположном случае она несовместна. Таким образом, теореме Кронекера—Капелли можно дать такую формулировку: *система линейных уравнений* (1) *тогда и только тогда совместна, если все ее характеристические определители равны нулю.*

Предположим теперь, что система (1) совместна. Теорема Кронекера—Капелли, при помощи которой мы устанавливаем совместность этой системы, утверждает существование решения; она не дает, однако, никакого способа для практического разыскания всех решений системы. К этой задаче мы сейчас переходим.

Пусть матрица A имеет ранг r. Как доказано в предшествующем параграфе, r равно максимальному числу линейно независимых строк матрицы A. Пусть, для определенности, первые r строк матрицы A линейно независимы, а каждая из остальных будет их линейной комбинацией. Тогда первые r строк матрицы $\bar{A}$ также будут линейно независимыми: всякая линейная зависимость между ними была бы линейной зависимостью и между первыми r строками матрицы A (вспомнить определение сложения векторов!). Из совпадения рангов матриц A и $\bar{A}$ следует, далее, что первые r строк матрицы $\bar{A}$ составляют в ней максимальную линейно независимую систему строк, т. е. всякая другая строка этой матрицы будет их линейной комбинацией.

Отсюда следует, что всякое уравнение системы (1) можно представить как сумму первых r уравнений, взятых с некоторыми коэффициентами, а поэтому любое общее решение первых r уравнений

будет удовлетворять всем уравнениям системы (1). Достаточно, следовательно, найти все решения системы

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \ldots\ldots\ldots\ldots&\ldots\ldots\\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rn}x_n&=b_r. \end{aligned}\right\} \qquad (2)$$

Так как строки из коэффициентов при неизвестных в уравнениях (2) линейно независимы, т. е. матрица из коэффициентов имеет ранг r, то $r \leqslant n$ и, кроме того, хотя бы один из миноров r-го порядка этой матрицы отличен от нуля. Если $r=n$, то (2) будет системой с равным числом уравнений и неизвестных и с отличным от нуля определителем, т. е. она, а потому и система (1), обладает единственным решением, а именно — вычисляемым по правилу Крамера.

Пусть теперь $r<n$ и пусть, для определенности, отличен от нуля минор r-го порядка, составленный из коэффициентов при первых r неизвестных. Перенесем в каждом из уравнений (2) в правую часть все члены с неизвестными $x_{r+1}, \ldots, x_n$ и выберем для этих неизвестных некоторые значения $c_{r+1}, \ldots, c_n$. Мы получим систему r уравнений

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1r}x_r&=b_1-a_{1,\,r+1}c_{r+1}-\ldots-a_{1n}c_n,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2r}x_r&=b_2-a_{2,\,r+1}c_{r+1}-\ldots-a_{2n}c_n,\\ \ldots\ldots\ldots\ldots&\ldots\ldots\ldots\ldots\\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rr}x_r&=b_r-a_{r,\,r+1}c_{r+1}-\ldots-a_{rn}c_n \end{aligned}\right\} \qquad (3)$$

относительно r неизвестных $x_1, x_2, \ldots, x_r$. К этой системе применимо правило Крамера, и поэтому она обладает единственным решением $c_1, c_2, \ldots, c_r$; очевидно, что система чисел $c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n$ будет служить решением системы (2). Так как значения $c_{r+1}, \ldots, c_n$ для неизвестных $x_{r+1}, \ldots, x_n$, называемых *свободными неизвестными*, мы могли выбирать произвольным образом, то этим путем будет получено бесконечно много различных решений системы (2).

С другой стороны, всякое решение системы (2) может быть получено указанным путем: если дано некоторое решение $c_1, c_2, \ldots, c_n$ системы (2), то в качестве значений для свободных неизвестных берем числа $c_{r+1}, \ldots, c_n$. Тогда числа $c_1, c_2, \ldots, c_r$ будут удовлетворять системе (3) и поэтому будут составлять то единственное решение этой системы, которое вычисляется по правилу Крамера.

Все сказанное выше объединяется в виде следующего правила решения произвольной системы линейных уравнений:

Пусть дана совместная система линейных уравнений (1) *и пусть матрица из коэффициентов A имеет ранг r. Выбираем в A r линейно независимых строк и оставляем в системе* (1)

лишь уравнения, коэффициенты которых вошли в выбранные строки. В этих уравнениях оставляем в левых частях такие r неизвестных, что определитель из коэффициентов при них отличен от нуля, а остальные неизвестные объявляем свободными и переносим в правые части уравнений. Давая свободным неизвестным произвольные числовые значения и вычисляя значения остальных неизвестных по правилу Крамера, мы получим все решения системы (1).

Дополнительно еще раз формулируем следующий полученный нами результат:

Совместная система (1) *тогда и только тогда обладает единственным решением, если ранг матрицы A равен числу неизвестных.*

Примеры. 1. Решить систему

$$\left.\begin{aligned} 5x_1 - x_2 + 2x_3 + x_4 &= 7, \\ 2x_1 + x_2 + 4x_3 - 2x_4 &= 1, \\ x_1 - 3x_2 - 6x_3 + 5x_4 &= 0. \end{aligned}\right\}$$

Ранг матрицы из коэффициентов равен двум: минор второго порядка, стоящий в левом верхнем углу этой матрицы, отличен от нуля, но оба минора третьего порядка, его окаймляющие, равны нулю. Ранг расширенной матрицы равен трем, так как

$$\begin{vmatrix} 5 & -1 & 7 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{vmatrix} = -35 \neq 0.$$

Отсюда следует, что система несовместна.

2. Решить систему

$$\left.\begin{aligned} 7x_1 + 3x_2 &= 2, \\ x_1 - 2x_2 &= -3, \\ 4x_1 + 9x_2 &= 11. \end{aligned}\right\}$$

Ранг матрицы из коэффициентов равен двум, т. е. равен числу неизвестных; ранг расширенной матрицы также равен двум. Таким образом, система совместна и обладает единственным решением. Левые части первых двух уравнений линейно независимы; решая систему этих двух уравнений, мы получим для неизвестных значения

$$x_1 = -\frac{5}{17}, \quad x_2 = \frac{23}{17}.$$

Легко видеть, что это решение удовлетворяет и третьему уравнению.

3. Решить систему

$$\left.\begin{aligned} x_1 + x_2 - 2x_3 - x_4 + x_5 &= 1, \\ 3x_1 - x_2 + x_3 + 4x_4 + 3x_5 &= 4, \\ x_1 + 5x_2 - 9x_3 - 8x_4 + x_5 &= 0. \end{aligned}\right\}$$

Система совместная, так как ранг расширенной матрицы, как и ранг матрицы из коэффициентов, равен двум. Левые части первого и третьего уравнений линейно независимы, так как коэффициенты при неизвестных x_1

и x_2 составляют отличный от нуля минор второго порядка. Решаем систему из этих двух уравнений, причем неизвестные x_3, x_4, x_5 считаем свободными, переносим в правые части уравнений и предполагаем, что им уже приданы некоторые числовые значения. Мы получим, применяя правило Крамера:

$$x_1 = \frac{5}{4} + \frac{1}{4}x_3 - \frac{3}{4}x_4 - x_5,$$

$$x_2 = -\frac{1}{4} + \frac{7}{4}x_3 + \frac{7}{4}x_4.$$

Эти равенства определяют *общее решение* заданной системы: давая в них свободным неизвестным произвольные числовые значения, мы получим все решения нашей системы. Так, решениями нашей системы будут, например, векторы $(2, 5, 3, 0, 0)$, $(3, 5, 2, 1, -2)$, $\left(0, -\frac{1}{4}, -1, 1, \frac{1}{4}\right)$ и т. д. С другой стороны, подставляя выражения для x_1 и x_2 из общего решения в любое из уравнений системы, например во второе, ранее исключенное из рассмотрения, мы получим тождество.

4. Решить систему

$$\left.\begin{aligned} 4x_1 + x_2 - 2x_3 + x_4 &= 3, \\ x_1 - 2x_2 - x_3 + 2x_4 &= 2, \\ 2x_1 + 5x_2 \qquad - x_4 &= -1, \\ 3x_1 + 3x_2 - x_3 - 3x_4 &= 1. \end{aligned}\right\}$$

Хотя число уравнений равно числу неизвестных, но определитель системы равен нулю и поэтому правило Крамера неприменимо. Ранг матрицы из коэффициентов равен трем — в правом верхнем углу этой матрицы расположен отличный от нуля минор третьего порядка. Ранг расширенной матрицы также равен трем, т. е. система совместна. Рассматривая лишь первые три уравнения и считая неизвестное x_1 свободным, мы получим общее решение в виде

$$x_2 = -\frac{1}{5} - \frac{2}{5}x_1, \quad x_3 = -\frac{8}{5} + \frac{9}{5}x_1, \quad x_4 = 0.$$

5. Пусть дана система, состоящая из $n+1$ уравнений относительно n неизвестных. Расширенная матрица $\bar{A}$ этой системы будет квадратной порядка $n+1$. Если наша система совместна, то, по теореме Кронекера—Капелли, определитель матрицы $\bar{A}$ должен быть равным нулю.

Так, пусть дана система

$$\left.\begin{aligned} x_1 - 8x_2 &= 3, \\ 2x_1 + x_2 &= 1, \\ 4x_1 + 7x_2 &= -4. \end{aligned}\right\}$$

Определитель из коэффициентов и свободных членов этих уравнений отличен от нуля:

$$\begin{vmatrix} 1 & -8 & 3 \\ 2 & 1 & 1 \\ 4 & 7 & -4 \end{vmatrix} = -77,$$

поэтому система несовместна.

Обратное утверждение не будет, вообще говоря, справедливым: из равенства нулю определителя матрицы $\bar{A}$ не следует совпадение рангов матриц A и $\bar{A}$.

§ 12. Системы линейных однородных уравнений

Применим результаты предшествующего параграфа к случаю *системы линейных однородных уравнений:*

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=0. \end{array}\right\} \quad (1)$$

Из теоремы Кронекера—Капелли вытекает, что эта система всегда совместна, так как добавление столбца из нулей не может повысить ранга матрицы. Это, впрочем, видно и непосредственно — система (1) заведомо обладает *нулевым решением* $(0, 0, \ldots, 0)$.

Пусть матрица A из коэффициентов системы (1) имеет ранг r. *Если $r=n$, то нулевое решение будет единственным решением системы* (1); *при $r<n$ система обладает также решениями, отличными от нулевого,* и для разыскания всех этих решений применяется тот же прием, как выше в случае произвольной системы уравнений. В частности, *система n линейных однородных уравнений с n неизвестными тогда и только тогда обладает решениями, отличными от нулевого, если определитель этой системы равен нулю*[1]. В самом деле, равенство нулю этого определителя равносильно утверждению, что ранг матрицы A меньше n. С другой стороны, *если в системе однородных уравнений число уравнений меньше числа неизвестных, то система непременно обладает решениями, отличными от нулевого,* так как ранг в этом случае не может быть равным числу неизвестных; этот результат уже был получен в § 1 из других соображений.

Рассмотрим, в частности, случай системы, состоящей из $n-1$ однородных уравнений относительно n неизвестных, причем предположим, что левые части этих уравнений между собой линейно независимы. Пусть

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ a_{21} & a_{22} & \ldots a_{2n} \\ \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & \ldots a_{n-1,n} \end{pmatrix}$$

— матрица из коэффициентов этой системы; через M_i обозначим минор $(n-1)$-го порядка, получающийся после вычеркивания из матрицы A ее i-го столбца, $i=1, 2, \ldots, n$. *Тогда одним из решений нашей системы будет система чисел*

$$M_1, \ -M_2, \ M_3, \ -M_4, \ \ldots, \ (-1)^{n-1}M_n, \quad (2)$$

а всякое другое решение ему пропорционально.

---

[1]) Одна половина этого утверждения уже была доказана в § 7.

Доказательство. Так как, по условию, ранг матрицы A равен $n-1$, то один из миноров M_i должен быть отличен от нуля; пусть это будет M_n. Полагаем в нашей системе неизвестное x_n свободным и переносим его в правую часть каждого из уравнений, после чего получим

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,\,n-1}x_{n-1} &= -a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2,\,n-1}x_{n-1} &= -a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n-1,\,1}x_1 + a_{n-1,\,2}x_2 + \ldots + a_{n-1,\,n-1}x_{n-1} &= -a_{n-1,\,n}x_n. \end{aligned}$$

Применяя затем правило Крамера, мы получим общее решение заданной системы уравнений, которому после легких преобразований можно придать вид

$$x_i = (-1)^{n-i}\frac{M_i}{M_n}x_n, \quad i = 1, 2, \ldots, n-1. \tag{3}$$

Положив $x_n = (-1)^{n-1}M_n$, мы получим: $x_i = (-1)^{2n-i-1}M_i$, $i = 1, 2, \ldots, n-1$, или, так как разность $(2n-i-1)-(i-1) = 2n-2i$ есть четное число, $x_i = (-1)^{i-1}M_i$, т. е. система чисел (2) действительно будет решением нашей системы уравнений. Любое другое решение этой системы получается из формул (3) при другом числовом значении неизвестного x_n, и поэтому оно пропорционально решению (2). Понятно, что рассматриваемое утверждение справедливо и в том случае, когда $M_n = 0$, но один из миноров M_i, $1 \leqslant i \leqslant n-1$, отличен от нуля.

Решения системы линейных однородных уравнений обладают следующими свойствами. Если вектор $\beta = (b_1, b_2, \ldots, b_n)$ является решением системы (1), то при любом числе k вектор $k\beta = (kb_1, kb_2, \ldots, kb_n)$ также будет решением этой системы, что проверяется непосредственной подстановкой в любое из уравнений (1). Если, далее, вектор $\gamma = (c_1, c_2, \ldots, c_n)$ — еще одно решение системы (1), то для этой системы служит решением и вектор $\beta + \gamma = (b_1 + c_1, b_2 + c_2, \ldots, b_n + c_n)$:

$$\sum_{j=1}^{n} a_{ij}(b_j + c_j) = \sum_{j=1}^{n} a_{ij}b_j + \sum_{j=1}^{n} a_{ij}c_j = 0, \quad i = 1, 2, \ldots, s.$$

Поэтому вообще *всякая линейная комбинация решений однородной системы* (1) *будет сама решением этой системы*. Заметим, что в случае *неоднородной* системы, т. е. системы линейных уравнений, свободные члены которых не все равны нулю, соответствующее утверждение не имеет места: ни сумма двух решений системы неоднородных уравнений, ни произведение решения этой системы на число не будут уже служить решениями для этой системы.

Мы знаем из § 9, что всякая система n-мерных векторов, состоящая более чем из n векторов, будет линейно зависимой. Отсюда следует, что из числа решений однородной системы (1), являющихся, как мы знаем, n-мерными векторами, можно выбрать конечную максимальную линейно независимую систему, максимальную в том смысле, что всякое другое решение системы (1) будет линейной

комбинацией решений, входящих в эту выбранную систему. Всякая максимальная линейно независимая система решений однородной системы уравнений (1) называется ее *фундаментальной системой решений.*

Еще раз подчеркнем, что *n-мерный вектор тогда и только тогда будет решением системы* (1), *если он является линейной комбинацией векторов, составляющих данную фундаментальную систему.*

Понятно, что фундаментальная система будет существовать лишь в том случае, если система (1) обладает ненулевыми решениями, т. е. если ранг ее матрицы из коэффициентов меньше числа неизвестных. При этом система (1) может обладать многими различными фундаментальными системами решений. Все эти системы эквивалентны, однако, между собой, так как каждый вектор всякой из этих систем линейно выражается через любую другую систему, и поэтому системы *состоят из одного и того же числа решений.*

Справедлива следующая теорема:

Если ранг r матрицы из коэффициентов системы линейных однородных уравнений (1) *меньше числа неизвестных n, то всякая фундаментальная система решений системы* (1) *состоит из $n-r$ решений.*

Для доказательства заметим, что $n-r$ является числом свободных неизвестных в системе (1); пусть свободными будут неизвестные $x_{r+1}, x_{r+2}, \ldots, x_n$. Рассмотрим произвольный отличный от нуля определитель d порядка $n-r$, который запишем в следующем виде:

$$d=\begin{vmatrix} c_{1,\,r+1}, & c_{1,\,r+2}, & \ldots, & c_{1n} \\ c_{2,\,r+1}, & c_{2,\,r+2}, & \ldots, & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n-r,\,r+1}, & c_{n-r,\,r+2}, & \ldots, & c_{n-r,\,n} \end{vmatrix}.$$

Беря элементы i-й строки этого определителя, $1 \leqslant i \leqslant n-r$, в качестве значений для свободных неизвестных, мы, как известно, получим однозначно определенные значения для неизвестных $x_1, x_2, \ldots, x_r$, т. е. придем к вполне определенному решению системы уравнений (1); запишем это решение в виде вектора

$$\alpha_i=(c_{i1}, c_{i2}, \ldots, c_{ir}, c_{i,\,r+1}, c_{i,\,r+2}, \ldots, c_{in}).$$

Полученная нами система векторов $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ служит для системы уравнений (1) фундаментальной системой решений. В самом деле, эта система векторов линейно независима, так как матрица, составленная из этих векторов как из строк, содержит отличный от нуля минор d порядка $n-r$. С другой стороны, пусть

$$\beta=(b_1, b_2, \ldots, b_r, b_{r+1}, b_{r+2}, \ldots, b_n)$$

будет произвольное решение системы уравнений (1). Докажем, что вектор β линейно выражается через векторы $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$.

Обозначим через α_i', $i=1, 2, \ldots, n-r$, i-ю строку определителя d, рассматриваемую как $(n-r)$-мерный вектор. Положим, далее,

$$\beta' = (b_{r+1}, b_{r+2}, \ldots, b_n).$$

Векторы α_i', $i=1, 2, \ldots, n-r$, линейно независимы, так как $d \neq 0$. Однако система $(n-r)$-мерных векторов

$$\alpha_1', \alpha_2', \ldots, \alpha_{n-r}', \beta'$$

линейно зависима, так как в ней число векторов больше их размерности. Существуют, следовательно, такие числа $k_1, k_2, \ldots, k_{n-r}$, что

$$\beta' = k_1\alpha_1' + k_2\alpha_2' + \ldots + k_{n-r}\alpha_{n-r}'. \quad (4)$$

Рассмотрим теперь n-мерный вектор

$$\delta = k_1\alpha_1 + k_1\alpha_2 + \ldots + k_{n-r}\alpha_{n-r} - \beta.$$

Вектор δ, являясь линейной комбинацией решений системы однородных уравнений (1), сам будет решением этой системы. Из (4) следует, что в решении δ значения для всех свободных неизвестных равны нулю. Однако то единственное решение системы уравнений (1), которое получается при равных нулю значениях для свободных неизвестных, будет нулевым решением. Таким образом, $\delta = 0$, т. е.

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \ldots + k_{n-r}\alpha_{n-r}.$$

Теорема доказана.

Заметим, что проведенное выше доказательство позволяет утверждать, что мы получим в с е фундаментальные системы решений системы однородных уравнений (1), беря в качестве d всевозможные отличные от нуля определители порядка $n-r$.

Пример. Дана система линейных однородных уравнений

$$\left.\begin{aligned} 3x_1 + x_2 - 8x_3 + 2x_4 + x_5 &= 0, \\ 2x_1 - 2x_2 - 3x_3 - 7x_4 + 2x_5 &= 0, \\ x_1 + 11x_2 - 12x_3 + 34x_4 - 5x_5 &= 0, \\ x_1 - 5x_2 + 2x_3 - 16x_4 + 3x_5 &= 0. \end{aligned}\right\}$$

Ранг матрицы из коэффициентов равен двум, число неизвестных равно пяти, поэтому всякая фундаментальная система решений этой системы уравнений состоит из трех решений. Решим систему, ограничиваясь первыми

двумя линейно независимыми уравнениями и считая x_3, x_4, x_5 свободными неизвестными. Мы получим общее решение в виде

$$x_1=\frac{19}{8}x_3+\frac{3}{8}x_4-\frac{1}{2}x_5,$$

$$x_2=\frac{7}{8}x_3-\frac{25}{8}x_4+\frac{1}{2}x_5.$$

Берем, далее, следующие три линейно независимых трехмерных вектора (1, 0, 0), (0, 1, 0), (0, 0, 1). Подставляя компоненты каждого из них в общее решение в качестве значений для свободных неизвестных и вычисляя значения для x_1 и x_2, мы получим следующую фундаментальиую систему решений заданной системы уравнений:

$$\alpha_1=\left(\frac{19}{8},\ \frac{7}{8},\ 1,\ 0,\ 0\right),\quad \alpha_2=\left(\frac{3}{8},\ -\frac{25}{8},\ 0,\ 1,\ 0\right),$$

$$\alpha_3=\left(-\frac{1}{2},\ \frac{1}{2},\ 0,\ 0,\ 1\right).$$

Мы закончим параграф рассмотрением связи, существующей между решениями неоднородных и однородных систем. Пусть дана система линейных неоднородных уравнений:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{array}\right\} \qquad (5)$$

Система линейных однородных уравнений:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=0, \end{array}\right\} \qquad (6)$$

полученная из системы (5) заменой свободных членов нулями, называется *приведенной системой* для системы (5). Между решениями систем (5) и (6) существует тесная связь, как показывают следующие две теоремы.

I. *Сумма любого решения системы* (5) *с любым решением приведенной системы* (6) *снова будет решением системы* (5).

В самом деле, пусть $c_1, c_2, \ldots, c_n$ — решение системы (5), $d_1, d_2, \ldots, d_n$ — решение системы (6). Берем любое из уравнений системы (5), например k-е, и вставляем в него вместо неизвестных числа $c_1+d_1, c_2+d_2, \ldots, c_n+d_n$. Мы получим:

$$\sum_{j=1}^{n}a_{kj}(c_j+d_j)=\sum_{j=1}^{n}a_{kj}c_j+\sum_{j=1}^{n}a_{kj}d_j=b_k+0=b_k.$$

II. *Разность любых двух решений системы* (5) *служит решением для приведенной системы* (6).

Действительно, пусть $c_1, c_2, \ldots, c_n$ и $c'_1, c'_2, \ldots, c'_n$ — решения системы (5). Берем любое из уравнений системы (6), например k-е, и подставляем в него вместо неизвестных числа

$$c_1 - c'_1,\ c_2 - c'_2,\ \ldots,\ c_n - c'_n.$$

Мы получим:

$$\sum_{j=1}^{n} a_{kj}(c_j - c'_j) = \sum_{j=1}^{n} a_{kj}c_j - \sum_{j=1}^{n} a_{kj}c'_j = b_k - b_k = 0.$$

Из этих теорем вытекает, что, *найдя одно решение системы линейных неоднородных уравнений* (5) *и складывая его с каждым из решений приведенной системы* (6), *мы получим все решения системы* (5).

ГЛАВА ТРЕТЬЯ

АЛГЕБРА МАТРИЦ

§ 13. Умножение матриц

В предшествующих главах понятие матрицы было использовано в качестве существенного вспомогательного орудия при изучении систем линейных уравнений. Многочисленные другие применения этого понятия сделали его предметом большой самостоятельной теории, во многих своих частях выходящей за пределы нашего курса. Мы займемся сейчас основами этой теории, начинающимися с того, что в множестве всех квадратных матриц данного порядка своеобразным, но вполне мотивированным способом определяются две алгебраические операции — сложение и умножение. Мы начнем с определения умножения матриц; сложение матриц будет введено в § 15.

Из курса аналитической геометрии известно, что при повороте осей прямоугольной системы координат на плоскости на угол α координаты точки преобразуются по следующим формулам:

$$x = x' \cos\alpha - y' \sin\alpha,$$
$$y = x' \sin\alpha + y' \cos\alpha,$$

где x, y — старые координаты точки, x', y' — ее новые координаты; таким образом, x и y выражаются через x' и y' линейно с некоторыми числовыми коэффициентами. Во многих других случаях также приходится встречаться с заменой неизвестных (или переменных), при которой старые неизвестные линейно выражаются через новые; такую замену неизвестных называют обычно их линейным преобразованием (или линейной подстановкой). Мы приходим, следовательно, к такому определению:

Линейным преобразованием неизвестных называется такой переход от системы n неизвестных $x_1, x_2, \ldots, x_n$ к системе n неизвестных $y_1, y_2, \ldots, y_n$, при котором старые неизвестные выражаются через новые линейно с некоторыми числовыми коэффициентами:

$$\left.\begin{array}{l} x_1 = a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n, \\ x_2 = a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n. \end{array}\right\} \quad (1)$$

Линейное преобразование (1) вполне определяется своей матрицей из коэффициентов

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

так как два линейных преобразования с одной и той же матрицей могут отличаться друг от друга лишь буквами, принятыми для обозначения неизвестных; мы будем считать, однако, что выбор этих обозначений целиком находится в нашем распоряжении. Обратно, задавая произвольную матрицу n-го порядка, мы сейчас же можем написать линейное преобразование, для которого эта матрица служит матрицей коэффициентов. Таким образом, между линейными преобразованиями n неизвестных и квадратными матрицами n-го порядка существует взаимно однозначное соответствие, а поэтому всякому понятию, связанному с линейными преобразованиями, и всякому свойству этих преобразований должно соответствовать аналогичное понятие или свойство, относящееся к матрицам.

Рассмотрим вопрос о последовательном выполнении двух линейных преобразований. Пусть вслед за линейным преобразованием (1) выполнено линейное преобразование

$$\left.\begin{aligned} y_1 &= b_{11}z_1 + b_{12}z_2 + \dots + b_{1n}z_n, \\ y_2 &= b_{21}z_1 + b_{22}z_2 + \dots + b_{2n}z_n, \\ &\dots\dots\dots\dots\dots\dots \\ y_n &= b_{n1}z_1 + b_{n2}z_2 + \dots + b_{nn}z_n, \end{aligned}\right\} \qquad (2)$$

переводящее систему неизвестных $y_1, y_2, \dots, y_n$ в систему $z_1, z_2, \dots, z_n$; матрицу этого преобразования обозначим через B. Подставляя в (1) выражения для $y_1, y_2, \dots, y_n$ из (2), мы придем к линейным выражениям для неизвестных $x_1, x_2, \dots, x_n$ через неизвестные $z_1, z_2, \dots, z_n$. Таким образом, *результат последовательного выполнения двух линейных преобразований неизвестных снова будет линейным преобразованием.*

Пример. Результатом последовательного выполнения линейных преобразований

$$x_1 = 3y_1 - y_2, \qquad y_1 = z_1 + z_2,$$
$$x_2 = y_1 + 5y_2, \qquad y_2 = 4z_1 + 2z_2$$

будет линейное преобразование

$$x_1 = 3(z_1 + z_2) - (4z_1 + 2z_2) = -z_1 + z_2,$$
$$x_2 = (z_1 + z_2) + 5(4z_1 + 2z_2) = 21z_1 + 11z_2.$$

Обозначим через C матрицу линейного преобразования, являющегося результатом последовательного выполнения преобразований (1) и (2), и найдем закон, по которому ее элементы c_{ik}, $i, k = 1, 2, \dots, n$,

выражаются через элементы матриц A и B. Коротко записывая преобразования (1) и (2) в виде

$$x_i = \sum_{j=1}^{n} a_{ij} y_j, \quad i = 1, 2, \ldots, n; \quad y_j = \sum_{k=1}^{n} b_{jk} z_k, \quad j = 1, 2, \ldots, n,$$

мы получим

$$x_i = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{n} b_{jk} z_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) z_k, \quad i = 1, 2, \ldots, n.$$

Таким образом, коэффициент при z_k в выражении для x_i, т. е. элемент c_{ik} матрицы C, имеет вид

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk}; \qquad (3)$$

элемент матрицы C, стоящий в i-й строке и k-м столбце, равен сумме произведений соответственных элементов i-й строки матрицы A и k-го столбца матрицы B.

Формула (3), дающая выражение элементов матрицы C через элементы матриц A и B, позволяет при заданных матрицах A и B сразу написать матрицу C, минуя рассмотрение линейных преобразований, соответствующих матрицам A и B. Этим путем всякой паре квадратных матриц n-го порядка ставится в соответствие однозначно определенная третья матрица. Можно сказать, что мы определили в множестве всех квадратных матриц n-го порядка алгебраическую операцию; она называется *умножением матриц*, а матрица C — *произведением* матрицы A на матрицу B:

$$C = AB.$$

Еще раз сформулируем связь между линейными преобразованиями и умножением матриц:

Линейное преобразование неизвестных, полученное в результате последовательного выполнения двух линейных преобразований с матрицами A и B, имеет своей матрицей коэффициентов матрицу AB.

Примеры.

1) $$\begin{pmatrix} 4 & 9 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 4\cdot 1 + 9\cdot(-2) & 4\cdot(-3) + 9\cdot 1 \\ (-1)\cdot 1 + 3\cdot(-2) & (-1)\cdot(-3) + 3\cdot 1 \end{pmatrix} = \begin{pmatrix} -14 & -3 \\ -7 & 6 \end{pmatrix}.$$

2) $$\begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} = \begin{pmatrix} -6 & 1 & 3 \\ 6 & 2 & 9 \\ -12 & -3 & 14 \end{pmatrix}.$$

3) $$\begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 51 & 16 \\ 8 & 3 \end{pmatrix}.$$

4) Найти результат последовательного выполнения линейных преобразований

$$\begin{aligned} x_1 &= 5y_1 - y_2 + 3y_3, \\ x_2 &= y_1 - 2y_2, \\ x_3 &= \phantom{5y_1 -{}} 7y_2 - y_3 \end{aligned}$$

и

$$\begin{aligned} y_1 &= 2z_1 \phantom{{}+z_2} + z_3, \\ y_2 &= \phantom{2z_1 +{}} z_2 - 5z_3, \\ y_3 &= \phantom{2z_1 +{}} 2z_2. \end{aligned}$$

Перемножая матрицы, получим:

$$\begin{pmatrix} 5 & -1 & 3 \\ 1 & -2 & 0 \\ 0 & 7 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -5 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 5 & 10 \\ 2 & -2 & 11 \\ 0 & 5 & -35 \end{pmatrix},$$

поэтому искомое линейное преобразование имеет вид:

$$\begin{aligned} x_1 &= 10z_1 + 5z_2 + 10z_3, \\ x_2 &= 2z_1 - 2z_2 + 11z_3, \\ x_3 &= \phantom{10z_1 +{}} 5z_2 - 35z_3. \end{aligned}$$

Возьмем один из рассмотренных сейчас примеров умножения матриц, например 2), и найдем произведение тех же матриц, но взятых в обратном порядке:

$$\begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} = \begin{pmatrix} -8 & 3 & -1 \\ 0 & 5 & 9 \\ 14 & -6 & 13 \end{pmatrix}.$$

Мы видим, что произведение матриц зависит от порядка множителей, т. е. *умножение матриц некоммутативно*. Этого, впрочем, следовало ожидать, уже хотя бы потому, что в определение матрицы C, данное выше при помощи формулы (3), матрицы A и B входят неравноправным образом: в A берутся строки, в B — столбцы.

Примеры неперестановочных матриц n-го порядка, т. е. матриц, произведение которых меняется при перестановке сомножителей, можно указать для всех n, начиная с $n=2$ (матрицы второго порядка в примере 1) неперестановочны). С другой стороны, две данные матрицы случайно могут оказаться перестановочными, как показывает следующий пример:

$$\begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} = \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} \cdot \begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

Умножение матриц ассоциативно; можно говорить, следовательно, об однозначно определенном произведении любого конечного числа матриц n-го порядка, взятых (ввиду некоммутативности умножения) в определенном порядке.

Доказательство. Пусть даны три произвольные матрицы n-го порядка A, B и C. Запишем их следующим сокращенным способом, указывающим общий вид их элементов: $A=(a_{ij})$, $B=(b_{ij})$, $C=(c_{ij})$. Введем, далее, следующие обозначения:

$$AB=U=(u_{ij}), \qquad BC=V=(v_{ij}),$$
$$(AB)\,C=S=(s_{ij}), \qquad A\,(BC)=T=(t_{ij}).$$

Нам нужно доказать справедливость равенства $(AB)\,C=A\,(BC)$, т. е. $S=T$. Однако

$$u_{il}=\sum_{k=1}^{n} a_{ik}b_{kl}, \quad v_{kj}=\sum_{l=1}^{n} b_{kl}c_{lj},$$

и поэтому, ввиду равенств $S=UC$, $T=AV$,

$$s_{ij}=\sum_{l=1}^{n} u_{il}c_{lj}=\sum_{l=1}^{n}\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj},$$
$$t_{ij}=\sum_{k=1}^{n} a_{ik}v_{kj}=\sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}b_{kl}c_{lj},$$

т. е. $s_{ij}=t_{ij}$ при $i,\ j=1,\ 2,\ \ldots,\ n$.

Дальнейшее изучение свойств умножения матриц требует привлечения их определителей, причем мы условимся для краткости обозначать определитель матрицы A через $|A|$. Если читатель в каждом из рассмотренных выше примеров подсчитает определители перемножаемых матриц и сравнит произведение этих определителей с определителем произведения заданных матриц, то обнаружит весьма любопытную закономерность, выражаемую следующей очень важной теоремой об умножении определителей:

Определитель произведения нескольких матриц n-го порядка равен произведению определителей этих матриц.

Достаточно доказать эту теорему для случая двух матриц. Пусть даны матрицы n-го порядка $A=(a_{ij})$ и $B=(b_{ij})$ и пусть $AB=C=(c_{ij})$. Построим следующий вспомогательный определитель Δ порядка $2n$: в его левом верхнем углу поставим матрицу A, в правом нижнем — матрицу B, весь правый верхний угол займем нулями и, наконец, по главной диагонали левого нижнего угла поставим число -1, заняв все остальные места также нулями. Определитель Δ имеет, следовательно, такой вид:

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 0 \\ -1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ 0 & -1 & \ldots & 0 & b_{21} & b_{22} & \ldots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & b_{n1} & b_{n2} & \ldots & b_{nn} \end{vmatrix}$$

Применение к определителю Δ теоремы Лапласа — разложение по первым n срокам — приводит к следующему равенству:

$$\Delta = |A| \cdot |B|. \tag{4}$$

Попытаемся, с другой стороны, так преобразовать определитель Δ, не меняя его значения, чтобы все элементы b_{ij}, $i, j = 1, 2, \ldots, n$, оказались замененными нулями. Для этой цели к $(n+1)$-му столбцу определителя Δ прибавим его первый столбец, умноженный на b_{11}, второй, умноженный на b_{21}, и т. д., наконец n-й столбец, умноженный на b_{n1}. Затем к $(n+2)$-му столбцу определителя Δ прибавим первый столбец, умноженный на b_{12}, второй, умноженный на b_{22}, и т. д. Вообще, к $(n+j)$-му столбцу определителя Δ, где $j = 1, 2, \ldots, n$, мы прибавим сумму первых n столбцов, взятых, соответственно, с коэффициентами $b_{1j}, b_{2j}, \ldots, b_{nj}$.

Легко видеть, что эти преобразования, не меняющие определителя, на самом деле приводят к замене всех элементов b_{ij} нулями. Одновременно вместо нулей, стоявших в правом верхнем углу определителя, появятся следующие числа: на пересечении i-й строки и $(n+j)$-го столбца определителя, $i, j = 1, 2, \ldots, n$, будет стоять теперь число $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$, равное, ввиду (3), элементу c_{ij} матрицы $C = AB$. Правый верхний угол определителя занимает теперь, следовательно, матрица C:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{vmatrix}.$$

Применим еще раз теорему Лапласа, разлагая определитель по последним n столбцам. Дополнительный минор для минора $|C|$ равен $(-1)^n$, а так как минор $|C|$ расположен в строках с номерами $1, 2, \ldots, n$ и в столбцах с номерами $n+1, n+2, \ldots, 2n$, причем

$$1 + 2 + \ldots + n + (n+1) + (n+2) + \ldots + 2n = 2n^2 + n,$$

то

$$\Delta = (-1)^{2n^2+n} (-1)^n |C| = (-1)^{2(n^2+n)} |C|$$

или, ввиду четности числа $2(n^2+n)$,

$$\Delta = |C|. \tag{5}$$

Из (4) и (5) вытекает, наконец, доказываемое равенство

$$|C|=|A|\cdot|B|.$$

Теорема об умножении определителей могла бы быть доказана и без использования теоремы Лапласа. Одно из таких доказательств читатель найдет в конце § 16.

§ 14. Обратная матрица

Квадратная матрица называется *вырожденной* (или *особенной*), если ее определитель равен нулю, и *невырожденной* (или *неособенной*) — в противоположном случае. Соответственно линейное преобразование неизвестных называется *вырожденным* или *невырожденным* в зависимости от того, будет ли равен нулю или отличен от нуля определитель из коэффициентов этого преобразования. Из теоремы, доказанной в конце предшествующего параграфа, вытекают следующие утверждения:

Произведение матриц, хотя бы одна из которых вырожденная, будет вырожденной матрицей.

Произведение любых невырожденных матриц само будет невырожденной матрицей.

Отсюда следует, ввиду связи, существующей между умножением матриц и последовательным выполнением линейных преобразований, такое утверждение: *результат последовательного выполнения нескольких линейных преобразований тогда и только тогда будет невырожденным преобразованием, если все заданные преобразования невырожденные.*

Роль единицы в умножении матриц играет единичная матрица

$$E=\begin{pmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ . & . & . & .\\ 0 & 0 & \dots & 1\end{pmatrix},$$

причем она перестановочна с любой матрицей A данного порядка,

$$AE=EA=A. \qquad (1)$$

Доказываются эти равенства или непосредственным применением правила умножения матриц, или же на основании замечания, что единичная матрица соответствует *тождественному* линейному преобразованию неизвестных

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= \quad y_2, \\ &\dots\dots \\ x_n &= \qquad y_n, \end{aligned}$$

выполнение которого до или после любого другого линейного преобразования, очевидно, не меняет этого последнего.

Заметим, что *матрица E является единственной матрицей, удовлетворяющей условию* (1) *при любой матрице A*. Действительно, если бы существовала еще матрица E' с этим же свойством, то мы имели бы

$$E'E = E', \quad E'E = E,$$

откуда $E' = E$.

Вопрос о существовании для данной матрицы A *обратной матрицы* оказывается более сложным. Ввиду некоммутативности умножения матриц мы будем говорить сейчас о *правой* обратной матрице, т. е. о такой матрице A^{-1}, что произведение матрицы A справа на эту матрицу дает единичную матрицу,

$$AA^{-1} = E. \tag{2}$$

Если матрица A вырожденная, то, если бы матрица A^{-1} существовала, произведение, стоящее в левой части равенства (2), было бы, как мы знаем, вырожденной матрицей, в то время как на самом деле матрица E, стоящая в правой части этого равенства, является невырожденной, так как ее определитель равен единице. Таким образом, вырожденная матрица не может иметь правой обратной матрицы. Такие же соображения показывают, что она не имеет и левой обратной и поэтому *для вырожденной матрицы обратная матрица вообще не существует*.

Переходя к случаю невырожденной матрицы, введем сначала следующее вспомогательное понятие. Пусть дана матрица n-го порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Матрица

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

составленная из алгебраических дополнений к элементам матрицы A, причем алгебраическое дополнение к элементу a_{ij} стоит на пересечении j-й строки и i-го столбца, называется *присоединенной* (или *взаимной*) матрицей к матрице A.

Найдем произведения AA^* и A^*A. Используя известную из § 6 формулу разложения определителя по строке или столбцу, а также теорему из § 7 о сумме произведений элементов любой строки

(столбца) определителя на алгебраические дополнения к соответственным элементам другой строки (столбца), и обозначая через d определитель матрицы A,

$$d = |A|,$$

мы получим следующие равенства:

$$AA^* = A^*A = \begin{pmatrix} d & 0 & \dots & 0 \\ 0 & d & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & d \end{pmatrix}. \qquad (3)$$

Отсюда вытекает, что *если матрица A невырожденная, то ее присоединенная матрица A^* также будет невырожденной, причем определитель d^* матрицы A^* равен $(n-1)$-й степени определителя d матрицы A.*

В самом деле, переходя от равенств (3) к равенству между определителями, мы получим

$$dd^* = d^n,$$

откуда ввиду $d \neq 0$

$$d^* = d^{n-1} \text{[1]}.$$

Теперь легко доказать существование обратной матрицы для всякой невырожденной матрицы A и найти ее вид. Заметим сначала, что если мы рассмотрим произведение двух матриц AB и все элементы одного из множителей, например B, разделим на одно и то же число d, то все элементы произведения AB также разделятся на это же число: для доказательства нужно лишь вспомнить определение умножения матриц. Таким образом, если

$$d = |A| \neq 0,$$

то из равенств (3) вытекает, что *обратной матрицей для A будет служить матрица, получающаяся из присоединенной матрицы A^* делением всех ее элементов на число d*:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{d} & \frac{A_{21}}{d} & \dots & \frac{A_{n1}}{d} \\ \frac{A_{12}}{d} & \frac{A_{22}}{d} & \dots & \frac{A_{n2}}{d} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{A_{1n}}{d} & \frac{A_{2n}}{d} & \dots & \frac{A_{nn}}{d} \end{pmatrix}.$$

---

[1]) Можно было бы доказать, что если матрица A вырожденная, то и ее присоединенная матрица A^* также вырожденная, причем имеет ранг, не превосходящий числа 1.

Действительно, из (3) вытекают равенства

$$AA^{-1}=A^{-1}A=E. \tag{4}$$

Еще раз подчеркнем, что в i-й строке матрицы A^{-1} стоят алгебраические дополнения к элементам i-го столбца определителя $|A|$, деленные на $d=|A|$.

Легко доказать, что *матрица A^{-1} является единственной матрицей, удовлетворяющей условию* (4) *для данной невырожденной матрицы A*. Действительно, если матрица C такова, что

$$AC=CA=E,$$

то

$$CAA^{-1}=C(AA^{-1})=CE=C,$$
$$CAA^{-1}=(CA)A^{-1}=EA^{-1}=A^{-1},$$

откуда $C=A^{-1}$.

Из (4) и теоремы об умножении определителей вытекает, что *определитель матрицы A^{-1} равен $\frac{1}{|A|}$, так что эта матрица также будет невырожденной; обратной для нее служит матрица A.*

Если теперь даны квадратные матрицы n-го порядка A и B, из которых A — невырожденная, а B — произвольная, то мы можем выполнить *правое и левое деления B на A*, т. е. решить матричные уравнения

$$AX=B, \quad YA=B. \tag{5}$$

Для этого, ввиду ассоциативности умножения матриц, достаточно положить

$$X=A^{-1}B, \quad Y=BA^{-1},$$

причем эти решения уравнений (5) будут, ввиду некоммутативности умножения матриц, в общем случае различными.

Примеры. 1) Дана матрица

$$A=\begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & 1 \\ 2 & -1 & 4 \end{pmatrix}.$$

Ее определитель $|A|=5$, поэтому обратная матрица A^{-1} существует, причем

$$A^{-1}=\begin{pmatrix} 1 & \frac{4}{5} & -\frac{1}{5} \\ 2 & \frac{12}{5} & -\frac{3}{5} \\ 0 & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

2) Даны матрицы

$$A=\begin{pmatrix}3 & 2\\ 4 & 3\end{pmatrix}, \quad B=\begin{pmatrix}-1 & 7\\ 3 & 5\end{pmatrix}.$$

Матрица A невырожденная, причем

$$A^{-1}=\begin{pmatrix}3 & -2\\ -4 & 3\end{pmatrix},$$

поэтому решениями уравнений $AX=B$, $YA=B$ будут служить матрицы

$$X=\begin{pmatrix}3 & -2\\ -4 & 3\end{pmatrix}\cdot\begin{pmatrix}-1 & 7\\ 3 & 5\end{pmatrix}=\begin{pmatrix}-9 & 11\\ 13 & -13\end{pmatrix},$$

$$Y=\begin{pmatrix}-1 & 7\\ 3 & 5\end{pmatrix}\cdot\begin{pmatrix}3 & -2\\ -4 & 3\end{pmatrix}=\begin{pmatrix}-31 & 23\\ -11 & 9\end{pmatrix}.$$

Умножение прямоугольных матриц. Хотя умножение матриц определено в предшествующем параграфе лишь для квадратных матриц одинакового порядка, но его можно распространить и на случай прямоугольных матриц A и B, если только можно применить формулу (3) предшествующего параграфа, т. е. если всякая строка матрицы A содержит столько же элементов, сколько их во всяком столбце матрицы B. Иными словами, *можно говорить о произведении прямоугольных матриц A и B в том случае, если число столбцов матрицы A равно числу строк матрицы B, причем число строк матрицы AB равно числу строк матрицы A, число же столбцов матрицы AB равно числу столбцов матрицы B.*

Примеры.

1) $$\begin{pmatrix}5 & -1 & 3 & 1\\ 2 & 0 & -1 & 4\end{pmatrix}\cdot\begin{pmatrix}-1 & 3 & 0\\ -2 & 1 & 1\\ 3 & 0 & -2\\ 4 & 1 & 2\end{pmatrix}=\begin{pmatrix}10 & 15 & -5\\ 11 & 10 & 10\end{pmatrix}.$$

2) $$\begin{pmatrix}0 & -3 & 1\\ 2 & 1 & 5\\ -4 & 0 & -2\end{pmatrix}\cdot\begin{pmatrix}3\\ -2\\ 2\end{pmatrix}=\begin{pmatrix}8\\ 14\\ -16\end{pmatrix}.$$

3) $$(5 \;\; 1 \;\; 0 \;\; -3)\cdot\begin{pmatrix}2 & 0\\ 1 & -4\\ 3 & 1\\ 0 & -1\end{pmatrix}=(11 \;\; -1).$$

Умножение прямоугольных матриц можно связать с последовательным выполнением линейных преобразований неизвестных, если только в определении последних отказаться от предположения, что число неизвестных сохраняется при линейном преобразовании.

Легко проверить также, дословно повторяя доказательство, данное выше для случая квадратных матриц, что *закон ассоциа-*

тивности остается справедливым и для умножения прямоугольных матриц.

Мы воспользуемся сейчас умножением прямоугольных матриц и свойствами обратной матрицы для нового вывода правила Крамера, не требующего тех громоздких вычислений, какие проводились в § 7. Пусть дана система n линейных уравнений с n неизвестными:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \cdots\cdots\cdots\cdots&\cdots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=b_n,\end{aligned}\right\} \tag{6}$$

причем определитель этой системы отличен от нуля. Обозначим через A матрицу из коэффициентов системы (6); эта матрица невырожденная, так как, по предположению, $d=|A|\neq 0$. Обозначим далее через X столбец из неизвестных, через B — столбец из свободных членов системы (6), т. е.

$$X=\begin{pmatrix} x_1\\ x_2\\ \cdot\\ \cdot\\ \cdot\\ x_n\end{pmatrix},\quad B=\begin{pmatrix} b_1\\ b_2\\ \cdot\\ \cdot\\ \cdot\\ b_n\end{pmatrix}.$$

Произведение AX имеет смысл, так как число столбцов матрицы A равно числу строк матриц X, причем это произведение будет столбцом, составленным из левых частей уравнений системы (6). Таким образом, систему (6) можно записать в виде одного матричного уравнения

$$AX=B. \tag{7}$$

Умножая обе части уравнения (7) слева на матрицу A^{-1}, существование которой вытекает из невырожденности квадратной матрицы A, мы получим:

$$X=A^{-1}B. \tag{8}$$

Произведение, стоящее справа, будет матрицей из одного столбца; ее j-й элемент равен сумме произведений элементов j-й строки матрицы A^{-1} на соответственные элементы матрицы B, т. е. равен числу

$$\frac{A_{1j}}{d}b_1+\frac{A_{2j}}{d}b_2+\ldots+\frac{A_{nj}}{d}b_n=\frac{1}{d}(A_{1j}b_1+A_{2j}b_2+\ldots+A_{nj}b_n).$$

Скобка, стоящая справа, является, однако, разложением по j-му столбцу определителя d_j, получающегося заменой j-го столбца определителя d столбцом B. Таким образом, формулы (8) равносильны

формулам (3) из § 7, выражающим решение системы (6), получающееся по правилу Крамера.

Остается показать, что полученные значения неизвестных действительно составляют решение системы (6). Для этого достаточно выражение (8) подставить в матричное уравнение (7), что, очевидно, приводит к тождеству $B=B$.

Ранг произведения матриц. Теорема об умножении определителей не приводит в случае вырожденных матриц ни к какому высказыванию сверх того, что их произведение также будет вырожденным, хотя вырожденные квадратные матрицы можно еще различать по их рангам. Заметим, что не существует вполне определенной зависимости между рангами сомножителей и рангом произведения, как показывают следующие примеры:

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

в обоих случаях перемножаются матрицы ранга 1, однако в одном случае произведение имеет ранг 1, в другом — ранг 0. Справедлива, притом не только для квадратных, но и для прямоугольных матриц, лишь следующая теорема.

Ранг произведения матриц не выше ранга каждого из сомножителей.

Достаточно доказать эту теорему для случая двух множителей. Пусть даны матрицы A и B, для которых произведение AB имеет смысл; обозначим $AB=C$. Рассмотрим формулу (3) § 13, дающую выражение элементов матрицы C. Беря эту формулу для данного k и всех возможных i $(i=1, 2, \ldots)$, мы получаем, что k-й столбец матрицы C является суммой всех столбцов матрицы A, взятых с некоторыми коэффициентами (а именно с коэффициентами $b_{1k}, b_{2k}, \ldots$). Этим доказано, что система столбцов матрицы C линейно выражается через систему столбцов матрицы A, а поэтому, как показано в § 9, ранг первой системы меньше или равен рангу второй системы; иными словами, ранг матрицы C не больше ранга матрицы A. Так как, с другой стороны, из той же формулы (3) § 13 при данном i и всех k вытекает, что всякая i-я строка матрицы C является линейной комбинацией строк матрицы B, то аналогичными рассуждениями мы получим, что ранг C не выше ранга B.

Более точный результат имеет место для случая, когда один из множителей является невырожденной квадратной матрицей:

Ранг произведения произвольной матрицы A справа или слева на невырожденную квадратную матрицу Q равен рангу матрицы A.

Пусть, например,

$$AQ=C. \qquad (9)$$

Из предшествующей теоремы следует, что ранг матрицы C не выше ранга матрицы A. Умножая, однако, равенство (9) справа на Q^{-1}, мы придем к равенству

$$A = CQ^{-1},$$

а поэтому, снова на основании предшествующей теоремы, ранг A не выше ранга C. Сопоставление этих двух результатов доказывает совпадение рангов матриц A и C.

§ 15. Сложение матриц и умножение матрицы на число

Для квадратных матриц порядка n следующим образом определяется сложение:

Суммой $A+B$ двух квадратных матриц $A=(a_{ij})$ и $B=(b_{ij})$ порядка n называется матрица $C=(c_{ij})$, всякий элемент которой равен сумме соответственных элементов матриц A и B;

$$c_{ij} = a_{ij} + b_{ij}\ [1].$$

Определенное нами сложение матриц будет, очевидно, коммутативным и ассоциативным. Для него существует обратная операция — вычитание, причем разностью матриц A и B служит матрица, составленная из разностей соответственных элементов заданных матриц. Роль нуля играет при этом *нулевая матрица*, составленная сплошь из нулей; в дальнейшем эта матрица обозначается символом 0: нет серьезной опасности смешать нулевую матрицу с числом нуль.

Сложение квадратных матриц и их умножение, определенное в § 13, *связаны законами дистрибутивности.*

В самом деле, пусть даны три матрицы порядка n, $A=(a_{ij})$, $B=(b_{ij})$, $C=(c_{ij})$. Тогда для любых i и j имеет место очевидное равенство

$$\sum_{s=1}^{n} (a_{is} + b_{is})\, c_{sj} = \sum_{s=1}^{n} a_{is}c_{sj} + \sum_{s=1}^{n} b_{is}c_{sj}.$$

Однако левая часть этого равенства есть элемент, стоящий в i-й строке и j-м столбце матрицы $(A+B)\,C$, правая часть — элемент, стоящий на этом же месте в матрице $AC+BC$. Этим доказано равенство

$$(A+B)\,C = AC + BC.$$

Равенство $C(A+B) = CA + CB$ доказывается таким же путем — некоммутативность умножения матриц требует, понятно, доказательства этих обоих законов дистрибутивности.

---

[1] Можно было бы, конечно, и произведение матриц определить столь же естественным способом, перемножая соответственные элементы. Такое умножение, однако, в отличие от того, которое определено в § 13, не нашло бы никаких серьезных применений.

Введем следующее определение умножения матриц на число. *Произведением* kA *квадратной матрицы* $A=(a_{ij})$ *на число* k называется матрица $A'=(a'_{ij})$, получающаяся умножением на k всех элементов матрицы A:

$$a'_{ij}=ka_{ij}.$$

С одним примером такого умножения матрицы на число мы уже встречались в предшествующем параграфе: если матрица A невырожденная, причем $|A|=d$, то ее обратная матрица A^{-1} и присоединенная матрица A^* связаны равенством

$$A^{-1}=d^{-1}A^*.$$

Как мы знаем, всякую квадратную матрицу порядка n можно рассматривать как n^2-мерный вектор, причем это соответствие между матрицами и векторами взаимно однозначное. Определенные сейчас сложение матриц и умножение матрицы на число превращаются при этом в сложение векторов и умножение вектора на число. Таким образом, *совокупность квадратных матриц порядка* n *можно рассматривать как* n^2*-мерное векторное пространство.*

Отсюда вытекает справедливость следующих равенств (здесь A, B — матрицы порядка n; k, l — числа, 1 — число единица):

$$k(A+B)=kA+kB, \tag{1}$$

$$(k+l)A=kA+lA, \tag{2}$$

$$k(lA)=(kl)A, \tag{3}$$

$$1\cdot A=A. \tag{4}$$

Свойства (1) и (2) связывают умножение матрицы на число со сложением матриц. Существует, вместе с тем, очень важная связь между умножением матрицы на число и перемножением самих матриц, а именно:

$$(kA)B=A(kB)=k(AB), \tag{5}$$

т. е. *если в произведении матриц один из множителей умножается на число* k, *то и все произведение будет умножаться на* k.

В самом деле, пусть даны матрицы $A=(a_{ij})$ и $B=(b_{ij})$ и число k. Тогда для любых i и j будет:

$$\sum_{s=1}^{n}(ka_{is})b_{sj}=k\sum_{s=1}^{n}a_{is}b_{sj}.$$

Левая часть этого равенства есть, однако, элемент, стоящий в i-й строке и j-м столбце матрицы $(kA)B$, правая часть — элемент,

стоящий на этом же месте в матрице $k(AB)$. Этим доказано равенство

$$(kA)B = k(AB).$$

Равенство $A(kB) = k(AB)$ доказывается таким же путем.

Операция умножения матрицы на число позволяет ввести новый способ записи матриц. Обозначим через E_{ij} матрицу, в которой на пересечении i-й строки и j-го столбца стоит единица, а все остальные элементы равны нулю. Полагая $i = 1, 2, \ldots, n$ и $j = 1, 2, \ldots, n$, мы получим n^2 таких матриц E_{ij}, которые связаны, как легко проверить, следующей таблицей умножения:

$$E_{is}E_{sj} = E_{ij}, \quad E_{is}E_{tj} = 0 \quad \text{при } s \neq t.$$

Матрица kE_{ij} отличается от матрицы E_{ij} лишь тем, что в ней на пересечении i-й строки и j-го столбца стоит число k. Учитывая это и используя определение сложения матриц, мы получаем следующую запись для произвольной квадратной матрицы A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}E_{ij}, \tag{6}$$

причем матрица A обладает, очевидно, лишь единственной записью вида (6).

Матрица kE, где E — единичная матрица, имеет, по определению умножения матрицы на число, следующий вид:

$$kE = \begin{pmatrix} k & & & 0 \\ & k & & \\ & & \ddots & \\ 0 & & & k \end{pmatrix},$$

т. е. на главной диагонали стоит одно и то же число k, а все элементы вне этой диагонали равны нулю. Такие матрицы называются *скалярными*.

Определение сложения матриц приводит к равенству

$$kE + lE = (k + l)E. \tag{7}$$

С другой стороны, используя определение умножения матриц или же опираясь на равенство (5), получаем:

$$kE \cdot lE = (kl)E. \tag{8}$$

Умножение матрицы A на число k можно истолковать как умножение A на скалярную матрицу kE в смысле перемножения матриц. Действительно, по (5),

$$(kE)A = A(kE) = kA.$$

Отсюда вытекает также, что *всякая скалярная матрица перестановочна с любой матрицей A*. Очень важно, что скалярные матрицы являются единственными, обладающими этим свойством:

Если некоторая матрица $C=(c_{ij})$ n-го порядка перестановочна с любой матрицей этого же порядка, то матрица C скалярна.

В самом деле, положим $i \neq j$ и рассмотрим равные между собой, по условию, произведения CE_{ij} и $E_{ij}C$ (см. выше определение матрицы E_{ij}). Легко видеть, что все столбцы матрицы CE_{ij}, кроме j-го, состоят из нулей, а j-й столбец совпадает с i-м столбцом матрицы C; в частности, на пересечении i-й строки и j-го столбца матрицы CE_{ij} стоит элемент c_{ii}. Аналогично все строки матрицы $E_{ij}C$, кроме i-й, состоят из нулей, а i-я строка совпадает с j-й строкой матрицы C; на пересечении i-й строки и j-го столбца матрицы $E_{ij}C$ расположен элемент c_{jj}. Используя равенство $CE_{ij}=E_{ij}C$, мы получаем, что $c_{ii}=c_{jj}$ (как элементы, стоящие в равных между собой матрицах на одинаковых местах), т. е. все элементы главной диагонали матрицы C равны между собой. С другой стороны, на пересечении j-й строки и j-го столбца матрица CE_{ij} стоит элемент c_{ji}; но в матрице $E_{ij}C$ на этом же месте стоит нуль (ввиду $i \neq j$), и поэтому $c_{ji}=0$, т. е. всякий элемент матрицы C, расположенный вне главной диагонали, равен нулю. Теорема доказана.

§ 16*. Аксиоматическое построение теории определителей

Определитель n-го порядка является числом, однозначно определяемым данной квадратной матрицей n-го порядка. Определение этого понятия, приведенное в § 4, указывает правило, по которому определитель выражается через элементы заданной матрицы. Это конструктивное определение можно, однако, заменить аксиоматическим; можно, иными словами, среди свойств определителя, установленных в §§ 4 и 6, указать такие, что единственной функцией матрицы с действительными значениями, обладающей этими свойствами, будет ее определитель.

Простейшее определение такого рода состоит в использовании разложений определителя по строке. Рассматриваем квадратные матрицы любых порядков и предполагаем, что всякой такой матрице M поставлено в соответствие число d_M, причем выполняются следующие условия:

1) Если матрица M первого порядка, т. е. состоит из одного элемента a, то $d_M=a$.

2) Если первую строку матрицы n-го порядка M составляют элементы $a_{11}, a_{12}, \ldots, a_{1n}$ и если через, M_i $i=1, 2, \ldots, n$, обозначена матрица $(n-1)$-го порядка, остающаяся после вычеркивания из M первой строки и i-го столбца, то

$$d_M=a_{11}d_{M_1}-a_{12}d_{M_2}+a_{13}d_{M_3}-\ldots+(-1)^{n-1}a_{1n}d_{M_n}.$$

Тогда для всякой матрицы M число d_M равно определителю этой матрицы. Мы предоставляем читателю доказательство этого утверждения, проводящееся индукцией по n и использующее результаты § 6.

Много более интересны другие формы аксиоматического определения определителя, относящиеся к случаю лишь одного данного порядка n и имеющие в своей основе некоторые из установленных в § 4 простейших свойств определителя. Мы приступим сейчас к рассмотрению одного из таких определений.

Пусть всякой квадратной матрице M n-го порядка поставлено в соответствие число d_M, причем выполняются следующие условия:

I. *Если одна из строк матрицы M умножается на число k, то число d_M также умножается на k.*

II. *Число d_M не меняется, если к одной из строк матрицы M прибавляется другая строка этой матрицы.*

III. *Если E — единичная матрица, то $d_E=1$.*

Докажем, что *для любой матрицы M число d_M равно определителю этой матрицы.*

Выведем сначала из условий I—III некоторые свойства числа d_M, аналогичные соответствующим свойствам определителя.

(1) *Если одна из строк матрицы M состоит из нулей, то $d_M=0$.*

В самом деле, умножая строку, состоящую из нулей, на число 0, мы не меняем матрицу, но, ввиду условия I, число d_M приобретает множитель 0. Поэтому

$$d_M=0\cdot d_M=0.$$

(2) *Число d_M не меняется, если к i-й строке матрицы M прибавляется ее j-я строка, $j\neq i$, умноженная на число k.*

Если $k=0$, то все доказано. Если же $k\neq 0$, то умножаем j-ю строку на k и получаем матрицу M', для которой, ввиду условия I, $d_{M'}=kd_M$. Затем к i-й строке матрицы M' прибавляем ее j-ю строку и получаем матрицу M'', причем, ввиду условия II, $d_{M''}=d_{M'}$. Наконец, умножаем j-ю строку матрицы M'' на число k^{-1}. Мы приходим к матрице M''', которая в действительности получена из M преобразованием, указанным в формулировке доказываемого свойства, причем

$$d_{M'''}=k^{-1}d_{M''}=k^{-1}d_{M'}=k^{-1}\cdot kd_M=d_M.$$

(3) *Если строки матрицы M линейно зависимы, то $d_M=0$.* Действительно, если одна из строк, например i-я, будет линейной комбинацией других строк, то, применяя несколько раз преобразование (2), можно i-ю строку заменить строкой из нулей. Преобразование (2) не меняет числа d_M, а поэтому ввиду свойства (1) $d_M=0$.

(4) *Если i-я строка матрицы M является суммой двух векторов β и γ и если матрицы M' и M'' получены из матрицы M заменой ее i-й строки соответственно векторами β и γ, то*

$$d_M = d_{M'} + d_{M''}.$$

В самом деле, пусть S будет система всех строк матрицы M, кроме i-й. Если в S существует линейная зависимость, то линейно зависимы строки каждой из матриц M, M' и M'', а поэтому, по свойству (3), $d_M = d_{M'} = d_{M''} = 0$, откуда следует справедливость в этом случае доказываемого свойства. Если же система S, состоящая из $n-1$ вектора, линейно независима, то, как показывают результаты § 9, ее можно дополнить некоторым вектором α до максимальной линейно независимой системы векторов n-мерного векторного пространства. Через эту систему можно линейно выразить векторы β и γ. Пусть вектор α входит в эти выражения с коэффициентами k и, соответственно, l; в выражение для вектора $\beta+\gamma$, т. е. для i-й строки матрицы M, вектор α будет входить, следовательно, с коэффициентом $k+l$. Матрицы M, M' и M'' можно теперь преобразовать, вычитая из их i-х строк некоторые линейные комбинации других строк так, что их i-ми строками будут служить соответственно векторы $(k+l)\alpha$, $k\alpha$ и $l\alpha$. Поэтому, обозначая через M^0 матрицу, получающуюся из матрицы M заменой ее i-й строки вектором α, и учитывая свойства (2) и I, мы приходим к равенствам:

$$d_M = (k+l)\, d_{M^0}, \quad d_{M'} = k d_{M^0}, \quad d_{M''} = l d_{M^0}.$$

Этим свойство (4) доказано.

(5) *Если матрица $\overline{M}$ получена из матрицы M транспозицией двух строк, то $d_{\overline{M}} = -d_M$.*

Пусть, в самом деле, в матрице M нужно переставить строки с номерами i и j. Этого можно достичь цепочкой следующих преобразований: сначала к i-й строке матрицы M прибавляем ее j-ю строку и получаем матрицу M', причем, по условию II, $d_{M'} = d_M$. Затем из j-й строки матрицы M' вычитаем ее i-ю строку и приходим к матрице M'', для которой, ввиду свойства (2), будет $d_{M''} = d_{M'}$; j-я строка матрицы M'' будет отличаться знаком от i-й строки матрицы M. Прибавим теперь к i-й строке матрицы M'' ее j-ю строку. Для матрицы M''', которую мы получим этим преобразованием, будет, по условию II, $d_{M'''} = d_{M''}$, причем i-я строка этой матрицы совпадет с j-й строкой матрицы M. Умножая, наконец, j-ю строку матрицы M''' на число -1, мы придем к искомой матрице $\overline{M}$. Поэтому, ввиду условия I,

$$d_{\overline{M}} = -d_{M'''} = -d_M.$$

(6) *Если матрица M' получена из матрицы M перестановкой строк, причем i-й строкой матрицы M', $i=1, 2, \ldots, n$, служит α_i-я строка матрицы M, то*

$$d_{M'} = \pm d_M;$$

при этом знак плюс соответствует случаю, когда подстановка

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$$

четна, знак минус — случаю, когда она нечетна.

В самом деле, матрица M' может быть получена из матрицы M некоторым числом транспозиций двух строк, и поэтому можно воспользоваться свойством (5). Четность числа этих транспозиций определяет, как известно из § 3, четность указанной выше подстановки.

Рассмотрим теперь матрицы $M=(a_{ij})$, $N=(b_{ij})$ и их произведение $Q=MN$ в смысле § 13. Найдем число d_Q. Мы знаем, что всякая i-я строка матрицы Q является суммой всех строк матрицы N, взятых соответственно с коэффициентами $a_{i1}, a_{i2}, \ldots, a_{in}$ (см., например, § 14). Заменим все строки матрицы Q их указанными линейными выражениями через строки матрицы N и воспользуемся несколько раз свойством (4). Мы получим, что число d_Q будет равно сумме чисел d_T для всевозможных матриц T следующего вида: i-я строка матрицы T, $i=1, 2, \ldots, n$, равна α_i-й строке матрицы N, умноженной на число $a_{i\alpha_i}$. При этом ввиду свойства (3) можно исключить из рассмотрения все матрицы T, для которых существуют такие индексы i и j, $i \neq j$, что $\alpha_i = \alpha_j$; остаются, иными словами, лишь такие матрицы T, для которых индексы $\alpha_1, \alpha_2, \ldots, \alpha_n$ составляют перестановку чисел $1, 2, \ldots, n$. Ввиду свойств I и (6) число d_T для такой матрицы имеет вид

$$d_T = \pm a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n} d_N,$$

где знак определяется четностью подстановки из индексов. Отсюда мы приходим к выражению для числа d_Q: после вынесения за скобки из всех слагаемых вида d_T общего множителя d_N в скобках остается, очевидно, определитель $|M|$ матрицы M в смысле конструктивного определения, данного в § 4, т. е.

$$d_Q = |M| \cdot d_N. \qquad (*)$$

Если мы возьмем теперь в качестве матрицы N единичную матрицу E, то будет $Q=M$ и, по свойству III, $d_N = d_E = 1$, т. е. *для любой матрицы M имеет место равенство*

$$d_M = |M|,$$

что и требовалось доказать. Одновременно еще раз, притом без использования теоремы Лапласа, доказана

теорема об умножении определителей: для этого достаточно в равенстве (*) заменить числа d_Q и d_N определителями соответствующих матриц.

Закончим эти аксиоматические рассмотрения доказательством независимости условий I—III, т. е. доказательством того, что ни одно из этих условий не является следствием двух других.

Для доказательства независимости условия III положим, что $d_M = 0$ для всякой матрицы M n-го порядка. Условия I и II будут, очевидно, выполняться, условие же III нарушается.

Для доказательства независимости условия II положим, что для всякой матрицы M число d_M равно произведению элементов, стоящих на главной диагонали этой матрицы. Условия I и III выполняются, а условие II уже не имеет места.

Наконец, для доказательства независимости условия I положим, что $d_M = 1$ для всякой матрицы M. Условия II и III будут при этом выполняться, а условие I нарушается.

ГЛАВА ЧЕТВЕРТАЯ

КОМПЛЕКСНЫЕ ЧИСЛА

§ 17. Система комплексных чисел

На протяжении курса элементарной алгебры несколько раз происходит обогащение запаса чисел. Школьник, приступающий к изучению алгебры, приносит из арифметики знакомство с положительными целыми и дробными числами. Алгебра начинается по существу с введения отрицательных чисел, т. е. с оформления первой среди важнейших числовых систем — системы *целых чисел*, состоящей из всех положительных и всех отрицательных целых чисел и нуля, и более широкой системы *рациональных чисел*, состоящей из всех целых чисел и всех дробных чисел, как положительных, так и отрицательных.

Дальнейшее расширение запаса чисел происходит тогда, когда в рассмотрение вводятся иррациональные числа. Система, состоящая из всех рациональных и всех иррациональных чисел, называется системой *действительных* (или *вещественных*) *чисел*. Строгое построение системы действительных чисел содержится обычно в университетском курсе математического анализа; для нас, однако, было достаточно в предшествующих главах и будет достаточно в дальнейшем того знакомства с действительными числами, каким обладает читатель, приступающий к изучению высшей алгебры.

Наконец, в самом конце курса элементарной алгебры система действительных чисел расширяется до системы *комплексных чисел*. Эта система чисел остается для читателя менее привычной, конечно, чем система действительных чисел, хотя на самом деле она обладает многими очень хорошими свойствами. В настоящей главе будет еще раз с необходимой полнотой изложена теория комплексных чисел.

Комплексные числа вводятся в связи со следующей задачей. Известно, что действительных чисел недостаточно для того, чтобы решить любое квадратное уравнение с действительными коэффициентами. Простейшее из квадратных уравнений, не имеющих корней среди действительных чисел, есть

$$x^2+1=0; \tag{1}$$

только это уравнение будет нас сейчас интересовать. Задача, стоящая перед нами, такова: *нужно расширить систему действительных чисел до такой системы чисел, в которой уравнение* (1) *уже обладало бы корнем.*

В качестве материала, из которого будет строиться эта новая система чисел, мы возьмем точки плоскости. Напомним, что изображение действительных чисел точками прямой линии (основанное на том, что мы получаем взаимно однозначное соответствие между множеством всех точек прямой и множеством всех действительных чисел, если при заданном начале координат и единице масштаба всякой точке прямой поставим в соответствие ее абсциссу) систематически используется во всех отделах математики и является столь привычным, что обычно мы не делаем различия между действительным числом и точкой, его изображающей.

Таким образом, *мы хотим определить систему чисел, изображающихся всеми точками плоскости.* До сих пор нам не приходилось складывать или перемножать точки плоскости, поэтому определение операций над точками мы имеем право выбирать, заботясь лишь о том, чтобы новая система чисел обладала всеми теми свойствами, ради которых мы ее создаем. Эти определения, особенно для произведения, покажутся в первый момент весьма искусственными. В гл. 10 будет показано, однако, что никакие другие определения операций, на первый взгляд даже более естественные, не привели бы нас к цели, т. е. к построению расширения системы действительных чисел, содержащего корень уравнения (1). Там же будет показано, что замена точек плоскости в этом построении любым другим материалом не привела бы к системе чисел, по своим алгебраическим свойствам отличающейся от той системы комплексных чисел, которая строится ниже.

Пусть на плоскости выбрана прямоугольная система координат. Условимся обозначать точки плоскости буквами α, β, γ, ... и записывать точку α с абсциссой a и ординатой b через (a, b), т. е., несколько отступая от того, что принято в аналитической геометрии, писать $\alpha=(a, b)$. Если даны точки $\alpha=(a, b)$ и $\beta=(c, d)$, то *суммой* этих точек мы будем называть точку с абсциссой $a+c$ и ординатой $b+d$, т. е.

$$(a, b)+(c, d)=(a+c, b+d); \tag{2}$$

произведением точек $\alpha=(a, b)$ и $\beta=(c, d)$ мы будем называть точку с абсциссой $ac-bd$ и ординатой $ad+bc$, т. е.

$$(a, b)(c, d)=(ac-bd, ad+bc). \tag{3}$$

Этим путем мы определили в множесте всех точек плоскости две алгебраические операции. Покажем, что *эти операции обладают всеми основными свойствами, какими обладают операции*

в системе действительных чисел или в системе рациональных чисел: они обе коммутативны и ассоциативны, связаны законом дистрибутивности и для них существуют обратные операции — вычитание и деление (кроме деления на нуль).

Коммутативность и ассоциативность сложения очевидны (точнее, вытекают из соответствующих свойств сложения действительных чисел), так как при сложении точек плоскости мы отдельно складываем их абсциссы и отдельно ординаты. Коммутативность умножения основана на том, что в определение произведения точки α и β входят симметричным образом. Ассоциативность умножения доказывают следующие равенства:

$$[(a,\ b)(c,\ d)](e,\ f)=(ac-bd,\ ad+bc)(e,\ f)=$$
$$=(ace-bde-adf-bcf,\ acf-bdf+ade+bce),$$
$$(a,\ b)[(c,\ d)(e,\ f)]=(a,\ b)(ce-df,\ cf+de)=$$
$$=(ace-adf-bcf-bde,\ acf+ade+bce-bdf).$$

Закон дистрибутивности вытекает из равенств

$$[(a,\ b)+(c,\ d)](e,\ f)=(a+c,\ b+d)(e,\ f)=$$
$$=(ae+ce-bf-df,\ af+cf+be+de),$$
$$(a,\ b)(e,f)+(c,\ d)(e,\ f)=(ae-bf,\ af+be)+(ce-df,\ cf+de)=$$
$$=(ae-bf+ce-df,\ af+be+cf+de).$$

Рассмотрим вопрос об обратных операциях. Если даны точки $\alpha=(a,\ b)$ и $\beta=(c,\ d)$, то их разностью будет такая точка $(x,\ y)$, что

$$(c,\ d)+(x,\ y)=(a,\ b).$$

Отсюда, ввиду (2),

$$c+x=a,\quad d+y=b.$$

Таким образом, *разностью* точек $\alpha=(a,\ b)$ и $\beta=(c,\ d)$ служит точка

$$\alpha-\beta=(a-c,\ b-d) \tag{4}$$

и эта разность однозначно определена. В частности, *нулем* будет служить начало координат (0, 0), а точкой, *противоположной* для точки $\alpha=(a,\ b)$, будет точка

$$-\alpha=(-a,\ -b). \tag{5}$$

Пусть, далее, даны точки $\alpha=(a,\ b)$ и $\beta=(c,\ d)$, причем точка β отлична от нуля, т. е. хотя бы одна из координат c, d не есть нуль и поэтому $c^2+d^2\neq 0$. Частным от деления α и β должна быть такая точка $(x,\ y)$, что $(c,\ d)(x,\ y)=(a,\ b)$. Отсюда, ввиду (3),

$$cx-dy=a,$$
$$dx+cy=b.$$

Решая эту систему уравнений, мы получим:

$$x=\frac{ac+bd}{c^2+d^2},\quad y=\frac{bc-ad}{c^2+d^2}.$$

Таким образом, при $\beta\neq0$ *частное* $\frac{\alpha}{\beta}$ существует и однозначно определено:

$$\frac{\alpha}{\beta}=\left(\frac{ac+bd}{c^2+d^2},\ \frac{bc-ad}{c^2+d^2}\right). \qquad (6)$$

Полагая здесь $\beta=\alpha$, мы получим, что *единицей* при нашем умножении точек служит точка (1, 0), лежащая на оси абсцисс на расстоянии 1 вправо от начала координат. Полагая, далее, в (6), что $\alpha=1=(1,0)$, мы получим, что при $\beta\neq0$ точкой, *обратной* для β, будет:

$$\beta^{-1}=\left(\frac{c}{c^2+d^2},\ \frac{-d}{c^2+d^2}\right). \qquad (7)$$

Таким образом, мы построили систему чисел, изображаемых точками плоскости, причем операции над этими числами определяются по формулам (2) и (3). Эта система чисел называется *системой комплексных чисел.*

Покажем, что *система комплексных чисел является расширением системы действительных чисел.* Для этой цели рассмотрим точки, лежащие на оси абсцисс, т. е. точки вида $(a, 0)$; ставя в соответствие точке $(a, 0)$ действительное число a, мы получаем, очевидно, взаимно однозначное соответствие между рассматриваемым множеством точек и множеством всех действительных чисел. Применение к этим точкам формул (2) и (3) дает равенства

$$(a, 0)+(b, 0)=(a+b, 0),$$
$$(a, 0)\cdot(b, 0)=(ab, 0),$$

т. е. точки $(a, 0)$ складываются и перемножаются друг с другом так же, как соответствующие действительные числа. Таким образом, *множество точек, лежащих на оси абсцисс, рассматриваемое как часть системы комплексных чисел, по своим алгебраическим свойствам ничем не отличается от системы действительных чисел, обычным способом изображенной точками прямой линии.* Это позволяет нам не различать в будущем точку $(a, 0)$ и действительное число a, т. е. всегда полагать $(a, 0)=a$. В частности, нуль (0, 0) и единица (1, 0) системы комплексных чисел оказываются обычными действительными числами 0 и 1.

Нам нужно теперь показать, что *среди комплексных чисел содержится корень уравнения* (1), т. е. такое число, квадрат которого равен действительному числу -1. Это будет, например, точка (0, 1), т. е. точка, лежащая на оси ординат на расстоянии 1

вверх от начала координат. Действительно, применяя (3), получаем:

$$(0, 1)\cdot(0, 1) = (-1, 0) = -1.$$

Условимся обозначать эту точку буквой i, так что $i^2 = -1$.

Покажем, наконец, что *для построенных нами комплексных чисел может быть получена их обычная запись.* Для этого найдем сначала произведение действительного числа b на точку i:

$$bi = (b, 0)\cdot(0, 1) = (0, b);$$

это будет, следовательно, точка, лежащая на оси ординат и имеющая ординату b, причем все точки оси ординат представимы в виде таких произведений. Если теперь (a, b) — произвольная точка, то ввиду равенства

$$(a, b) = (a, 0) + (0, b)$$

получаем:

$$(a, b) = a + bi,$$

т. е. мы действительно приходим к обычной записи комплексных чисел; произведение и сумму в выражении $a + bi$ следует понимать, конечно, в смысле операций, определенных в построенной нами системе комплексных чисел.

Теперь, когда комплексные числа нами уже построены, читатель без труда проверит, что *все содержание предшествующих глав книги* — и теория определителей, и теория систем линейных уравнений, и теория линейной зависимости векторов, и теория операций над матрицами — *без всяких ограничений переносится на тот случай, когда к рассмотрению допускаются любые комплексные числа, а не только числа действительные.*

В заключение заметим, что проведенное нами построение системы комплексных чисел подсказывает следующий вопрос: нельзя ли так определить сложение и умножение точек трехмерного пространства, чтобы совокупность этих точек стала системой чисел, содержащей в себе систему комплексных чисел или хотя бы систему действительных чисел? Этот вопрос выходит за рамки нашего курса, и мы лишь отметим, что ответ на него оказывается отрицательным.

С другой стороны, замечая, что сложение комплексных чисел, определенное выше, по существу совпадает со сложением векторов на плоскости, выходящих из начала координат (см. следующий параграф), естественно поставить такой вопрос: можно ли при некоторых n так определить умножение векторов в n-мерном действительном векторном пространстве, чтобы по отношению к этому умножению и обычному сложению векторов наше пространство оказалось числовой системой, содержащей в себе систему действительных чисел? Можно показать, что этого сделать нельзя, если требовать выполнения всех тех свойств операций, которые имеют место в системах рациональных, действительных и комплексных чисел.

Если же отказаться от коммутативности умножения, то такое построение возможно в четырехмерном пространстве; получающаяся система чисел называется *системой кватернионов*. Аналогичное построение возможно и в восьмимерном пространстве — получается так называемая *система чисел Кэли*. Здесь приходится отказываться, впрочем, не только от коммутативности умножения, но и от его ассоциативности, заменяя последнее одним более слабым требованием.

§ 18. Дальнейшее изучение комплексных чисел

В соответствии с исторически сложившимися традициями мы будем называть комплексное число i *мнимой единицей*, а числа вида bi — *чисто мнимыми числами*, хотя существование этих чисел не вызывает у нас сомнений и мы можем указать те точки плоскости — точки оси ординат, — которыми эти числа изображаются. В записи комплексного числа α в виде $\alpha = a + bi$ число a называется *действительной частью* числа α, а bi — его *мнимой частью*. Плоскость, точки которой отождествлены с комплексными числами по способу, изложенному в § 17, будет называться *комплексной плоскостью*. Ось абсцисс этой плоскости называется *действительной осью*, так как ее точки изображают действительные числа; соответственно ось ординат комплексной плоскости называется *мнимой осью*.

Сложение, умножение, вычитание и деление комплексных чисел, записанных в виде $a+bi$, производятся следующим образом, как вытекает из формул (2), (4), (3) и (6) предшествующего параграфа:

$$(a+bi)+(c+di)=(a+c)+(b+d)i;$$
$$(a+bi)-(c+di)=(a-c)+(b-d)i;$$
$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i;$$
$$\frac{a+bi}{c+di}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.$$

Мы можем сказать, что *при сложении комплексных чисел складываются отдельно их действительные части и отдельно их мнимые части*; аналогичное правило имеет место и для вычитания. Словесные выражения для формул умножения и деления были бы слишком громоздкими, и мы их не даем. Последнюю из этих формул нет необходимости запоминать; следует лишь помнить, что ее можно вывести, умножая числитель и знаменатель заданной дроби на число, отличающееся от знаменателя лишь знаком при мнимой части.

Действительно,

$$\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.$$

Примеры.

1) $(2+5i)+(1-7i)=(2+1)+(5-7)i=3-2i$;

2) $(3-9i)-(7+i)=(3-7)+(-9-1)i=-4-10i$;

3) $(1+2i)(3-i)=[1\cdot 3-2\cdot(-1)]+[1\cdot(-1)+2\cdot 3]i=5+5i$;

4) $\dfrac{23+i}{3+i}=\dfrac{(23+i)(3-i)}{(3+i)(3-i)}=\dfrac{70-20i}{10}=7-2i$.

Изображение комплексных чисел точками плоскости приводит к естественному желанию иметь геометрическое истолкование операций, определенных для комплексных чисел. Для сложения такое истолкование может быть получено без затруднений. Пусть даны числа $\alpha=a+bi$ и $\beta=c+di$. Соединяем соответствующие им точки (a, b) и (c, d) отрезками с началом координат и строим на этих отрезках, как на сторонах, параллелограмм (рис. 2). Четвертой вершиной этого параллелограмма будет, очевидно, точка

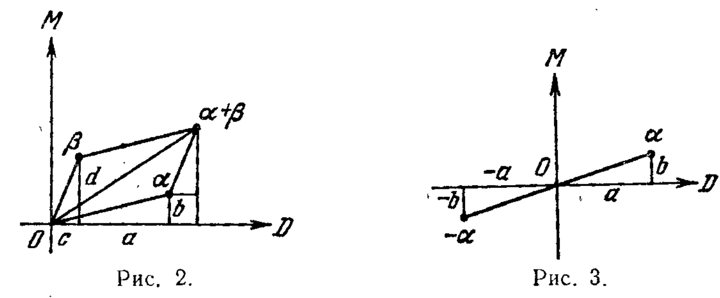


Рис. 2. Рис. 3.

$(a+c, b+d)$. Таким образом, *сложение комплексных чисел геометрически выполняется по правилу параллелограмма, т. е. по правилу сложения векторов, выходящих из начала координат.* Далее, *число, противоположное числу* $\alpha=a+bi$, *будет точкой комплексной плоскости, симметричной с точкой* α *относительно начала координат* (рис. 3). Отсюда без труда может быть получено геометрическое истолкование вычитания.

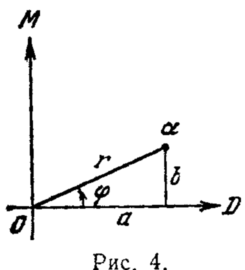


Рис. 4.

Геометрический смысл умножения и деления комплексных чисел станет ясным лишь после того, как мы введем для комплексных чисел новую запись, отличную от употреблявшейся нами до сих пор. Запись числа α в виде $\alpha=a+bi$ использует декартовы координаты точки, соответствующей этому числу. Положение точки на плоскости вполне определяется, однако, также заданием ее полярных координат: расстояния r от начала координат до точки и угла φ между положительным направлением оси абсцисс и направлением из начала координат на эту точку (рис. 4).

Число r является неотрицательным действительным числом, причем оно равно нулю лишь для точки 0. Для α, лежащего на действительной оси, т. е. являющегося действительным числом, число r будет абсолютной величиной α, поэтому и для любого комплексного числа α его иногда называют *абсолютной величиной* числа α; чаще, впрочем, число r называют *модулем* числа α. Обозначается оно через $|\alpha|$.

Угол φ будет называться *аргументом* числа α и обозначаться $\arg\alpha$[1]). Угол φ может принимать любые действительные значения, как положительные, так и отрицательные, причем положительные углы должны отсчитываться против часовой стрелки, однако, если углы отличаются друг от друга на 2π или число, кратное 2π, то соответствующие им точки плоскости совпадают.

Таким образом аргумент комплексного числа α имеет бесконечно много значений, отличающихся друг от друга на целые кратные числа 2π; из равенства двух комплексных чисел, заданных их модулями и аргументами, можно лишь заключить, следовательно, что аргументы отличаются на целое кратное числа 2π, в то время как модули равны. Аргумент не определен лишь для числа 0; это число вполне определяется, однако, равенством $|0|=0$.

Аргумент комплексного числа является естественным обобщением знака действительного числа. В самом деле, аргумент положительного действительного числа равен 0, аргумент отрицательного действительного числа равен π; на действительной оси из начала координат выходят лишь два направления и их можно различать двумя символами $+$ и $-$, тогда как на комплексной плоскости направлений, выходящих из точки 0, бесконечно много и различаются они уже углом, составляемым ими с положительным направлением действительной оси.

Между декартовыми и полярными координатами точки существует следующая связь, справедливая при любом расположении точек на плоскости:

$$a=r\cos\varphi,\quad b=r\sin\varphi. \tag{1}$$

Отсюда

$$r=+\sqrt{a^2+b^2}. \tag{2}$$

Применим формулы (1) к произвольному комплексному числу $\alpha=a+bi$:

$$\alpha=a+bi=r\cos\varphi+(r\sin\varphi)\,i,$$

или

$$\alpha=r(\cos\varphi+i\sin\varphi). \tag{3}$$

Обратно, пусть число $\alpha=a+bi$ допускает запись вида $\alpha=r_0(\cos\varphi_0+i\sin\varphi_0)$, где r_0 и φ_0 — некоторые действительные

---

[1]) Мы отказываемся, следовательно, от обычных названий полярных координат точки — полярный радиус и полярный угол.

числа, причем $r_0 \geqslant 0$. Тогда $r_0 \cos\varphi_0 = a$, $r_0 \sin\varphi_0 = b$, откуда $r_0 = +\sqrt{a^2+b^2}$, т. е., ввиду (2), $r_0 = |\alpha|$. Отсюда, используя (1), получаем: $\cos\varphi_0 = \cos\varphi$, $\sin\varphi_0 = \sin\varphi$, т. е. $\varphi_0 = \arg\alpha$. Таким образом, *всякое комплексное число α однозначным образом записывается в виде* (3), *где* $r = |\alpha|$, $\varphi = \arg\alpha$ (причем аргумент φ определен, конечно, лишь с точностью до слагаемых, кратных 2π). Эта запись числа α называется его *тригонометрической формой* и будет дальше весьма часто использоваться.

Числа

$$\alpha = 3\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right),\quad \beta = \cos\frac{19}{3}\pi + i\sin\frac{19}{3}\pi$$

и

$$\gamma = \sqrt{3}\left[\cos\left(-\frac{\pi}{7}\right) + i\sin\left(-\frac{\pi}{7}\right)\right]$$

заданы в тригонометрической форме; здесь $|\alpha| = 3$, $|\beta| = 1$, $|\gamma| = \sqrt{3}$; $\arg\alpha = \frac{\pi}{4}$, $\arg\beta = \frac{19}{3}\pi$, $\arg\gamma = -\frac{\pi}{7}$ $\left(\text{или } \arg\beta = \frac{\pi}{3},\ \arg\gamma = \frac{13}{7}\pi\right)$.

С другой стороны, комплексные числа

$$\alpha' = (-2)\left(\cos\frac{\pi}{5} + i\sin\frac{\pi}{5}\right),\quad \beta' = 3\left(\cos\frac{2}{3}\pi - i\sin\frac{2}{3}\pi\right),$$

$$\gamma' = 2\left(\cos\frac{\pi}{3} + i\sin\frac{3}{4}\pi\right),\quad \delta' = \sin\frac{3}{4}\pi + i\cos\frac{3}{4}\pi$$

даны не в тригонометрической форме, хотя их записи напоминают запись (3). В тригонометрической форме эти числа записываются так:

$$\alpha' = 2\left(\cos\frac{6}{5}\pi + i\sin\frac{6}{5}\pi\right),\quad \beta' = 3\left(\cos\frac{4}{3}\pi + i\sin\frac{4}{3}\pi\right),$$

$$\delta' = \cos\frac{7}{4}\pi + i\sin\frac{7}{4}\pi.$$

Разыскание тригонометрической формы числа γ' наталкивается на трудность, почти всегда встречающуюся при переходе от обычной записи комплексного числа к тригонометрической и обратно: невозможно, кроме немногих случаев, по заданным числовым значениям синуса и косинуса найти точно угол, а для заданного угла написать точные значения его синуса и косинуса.

Пусть комплексные числа α и β заданы в тригонометрической форме: $\alpha = r(\cos\varphi + i\sin\varphi)$, $\beta = r'(\cos\varphi' + i\sin\varphi')$. Перемножим эти числа:

$$\alpha\beta = [r(\cos\varphi + i\sin\varphi)]\cdot[r'(\cos\varphi' + i\sin\varphi')] =$$
$$= rr'(\cos\varphi\cos\varphi' + i\cos\varphi\sin\varphi' + i\sin\varphi\cos\varphi' - \sin\varphi\sin\varphi'),$$

или

$$\alpha\beta = rr'[\cos(\varphi+\varphi') + i\sin(\varphi+\varphi')]. \tag{4}$$

Мы получили запись произведения $\alpha\beta$ в тригонометрической форме, и поэтому $|\alpha\beta| = rr'$, или

$$|\alpha\beta| = |\alpha||\beta|, \tag{5}$$

т. е. *модуль произведения комплексных чисел равен произведению модулей сомножителей;* далее, $\arg(\alpha\beta) = \varphi + \varphi'$ или

$$\arg(\alpha\beta) = \arg\alpha + \arg\beta, \tag{6}$$

т. е. *аргумент произведения комплексных чисел равен сумме аргументов сомножителей*[1]). Эти правила распространяются, очевидно, на любое конечное число множителей. В применении к случаю действительных чисел формула (5) дает известное свойство абсолютных величин этих чисел, а (6) превращается, как легко проверить, в правило знаков при умножении действительных чисел.

Аналогичные правила имеют место и для частного. Действительно, пусть $\alpha = r(\cos\varphi + i\sin\varphi)$, $\beta = r'(\cos\varphi' + i\sin\varphi')$, причем $\beta \neq 0$, т. е. $r' \neq 0$. Тогда

$$\frac{\alpha}{\beta} = \frac{r(\cos\varphi + i\sin\varphi)}{r'(\cos\varphi' + i\sin\varphi')} = \frac{r(\cos\varphi + i\sin\varphi)(\cos\varphi' - i\sin\varphi')}{r'(\cos^2\varphi' + \sin^2\varphi')} =$$
$$= \frac{r}{r'}(\cos\varphi\cos\varphi' + i\sin\varphi\cos\varphi' - i\cos\varphi\sin\varphi' + \sin\varphi\sin\varphi'),$$

или

$$\frac{\alpha}{\beta} = \frac{r}{r'}[\cos(\varphi - \varphi') + i\sin(\varphi - \varphi')]. \tag{7}$$

Отсюда следует, что $\left|\frac{\alpha}{\beta}\right| = \frac{r}{r'}$ или

$$\left|\frac{\alpha}{\beta}\right| = \frac{|\alpha|}{|\beta|}, \tag{8}$$

т. е. *модуль частного двух комплексных чисел равен модулю делимого, деленному на модуль делителя;* далее, $\arg\left(\frac{\alpha}{\beta}\right) = \varphi - \varphi'$, или

$$\arg\left(\frac{\alpha}{\beta}\right) = \arg\alpha - \arg\beta, \tag{9}$$

т. е. *аргумент частного двух комплексных чисел получается вычитанием аргумента делителя из аргумента делимого.*

Геометрический смысл умножения и деления выясняется теперь без затруднений. Действительно, ввиду формул (5) и (6), мы получим точку, изображающую произведение числа α на число $\beta = r'(\cos\varphi' + i\sin\varphi')$, если вектор, идущий от 0 к α (рис. 5), повернем против часовой стрелки на угол $\varphi' = \arg\beta$, а затем

[1]) Подчеркнем, что равенство здесь понимается с точностью до слагаемого, кратного 2π.

растянем этот вектор в $r'=|\beta|$ раз (при $0\leqslant r'<1$ это будет, конечно, сжатием, а не растяжением). Далее, из (7) следует, что при $\alpha=r(\cos\varphi+i\sin\varphi)\neq 0$ будет

$$\alpha^{-1}=r^{-1}[\cos(-\varphi)+i\sin(-\varphi)], \qquad (10)$$

т. е. $|\alpha^{-1}|=|\alpha|^{-1}$, $\arg(\alpha^{-1})=-\arg\alpha$. Таким образом, мы получим точку α^{-1}, если от точки α перейдем к точке α', лежащей на расстоянии r^{-1} от нуля на той же полупрямой, выходящей из нуля,

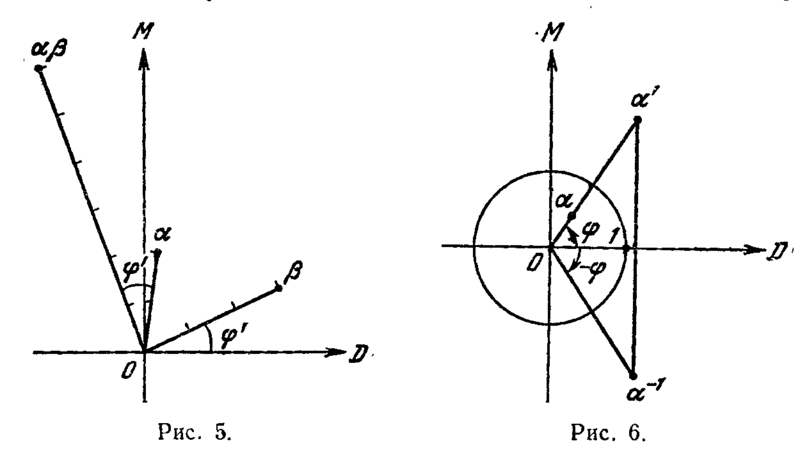


Рис. 5. Рис. 6.

что и точка α (рис. 6)[1], а затем перейдем к точке, симметричной с α' относительно действительной оси.

Сумму и разность комплексных чисел, заданных в тригонометрической форме, нельзя выразить формулами, подобными формулам (4) и (7). Для модуля суммы имеют место, однако, следующие важные неравенства:

$$|\alpha|-|\beta|\leqslant|\alpha+\beta|\leqslant|\alpha|+|\beta|, \qquad (11)$$

т. е. *модуль суммы двух комплексных чисел меньше или равен сумме модулей слагаемых, но больше или равен разности этих модулей.* Неравенства (11) вытекают из известной теоремы элементарной геометрии о сторонах треугольника ввиду того, что $|\alpha+\beta|$ равен, как мы знаем, диагонали параллелограмма со сторонами $|\alpha|$ и $|\beta|$. Специального рассмотрения, которое мы предоставляем читателю, требует, впрочем, случай, когда точки α, β и 0 лежат

[1]) Тогда и только тогда $|\alpha'|=|\alpha|$, если $|\alpha|=1$, т. е. если точка α лежит на окружности *единичного круга*. Если α лежит внутри единичного круга, то α' будет вне его, и обратно, причем этим путем мы получаем, очевидно, взаимно однозначное соответствие между всеми точками комплексной плоскости, лежащими вне единичного круга, и всеми точками, лежащими внутри этого круга и отличными от нуля.

на одной прямой; лишь в этом случае в формулах (11) могут достигаться равенства.

Из (11), ввиду $\alpha-\beta=\alpha+(-\beta)$ и

$$|-\beta|=|\beta| \qquad (12)$$

(это равенство следует хотя бы из геометрического толкования числа $-\beta$), вытекают также неравенства

$$|\alpha|-|\beta| \leqslant |\alpha-\beta| \leqslant |\alpha|+|\beta|, \qquad (13)$$

т. е. для модуля разности имеют место такие же неравенства, как и для модуля суммы.

Неравенства (11) можно было бы получить также следующим путем. Пусть $\alpha=r(\cos\varphi+i\sin\varphi)$, $\beta=r'(\cos\varphi'+i\sin\varphi')$ и пусть тригонометрическая форма числа $\alpha+\beta$ есть $\alpha+\beta=R(\cos\psi+i\sin\psi)$. Складывая отдельно действительные и отдельно мнимые части, получаем:

$$r\cos\varphi+r'\cos\varphi'=R\cos\psi,$$
$$r\sin\varphi+r'\sin\varphi'=R\sin\psi;$$

умножая обе части первого равенства на $\cos\psi$, обе части второго — на $\sin\psi$ и складывая, получаем:

$$r(\cos\varphi\cos\psi+\sin\varphi\sin\psi)+r'(\cos\varphi'\cos\psi+\sin\varphi'\sin\psi)=R(\cos^2\psi+\sin^2\psi),$$

т. е.

$$r\cos(\varphi-\psi)+r'\cos(\varphi'-\psi)=R.$$

Отсюда, так как косинус никогда не бывает больше единицы, следует неравенство $r+r' \geqslant R$, т. е. $|\alpha|+|\beta| \geqslant |\alpha+\beta|$. С другой стороны, $\alpha=(\alpha+\beta)-\beta=(\alpha+\beta)+(-\beta)$. Отсюда, по доказанному и в силу (12),

$$|\alpha| \leqslant |\alpha+\beta|+|-\beta|=|\alpha+\beta|+|\beta|,$$

откуда $|\alpha|-|\beta| \leqslant |\alpha+\beta|$.

Следует заметить, что для комплексных чисел понятия «больше» и «меньше» не могут быть разумно определены, так как эти числа, в отличие от действительных чисел, располагаются не на прямой линии, точки которой естественным образом упорядочены, а на плоскости. Поэтому сами комплексные числа (а не их модули) никогда нельзя соединять знаком неравенства.

Сопряженные числа. Пусть дано комплексное число $\alpha=a+bi$. Число $a-bi$, отличающееся от α лишь знаком при мнимой части, называется числом, *сопряженным с* α, и обозначается $\bar{\alpha}$.

Напомним, что при рассмотрении деления комплексных чисел мы прибегали к сопряженным числам, хотя и не вводили этого названия.

Числом, сопряженным с $\bar{\alpha}$, будет, очевидно, α, т. е. можно говорить о паре сопряженных чисел. Действительные числа, и только они, сопряжены сами себе.

Геометрически сопряженные числа являются точками, симметричными относительно действительной оси (рис. 7). Отсюда следуют равенства

$$|\bar{\alpha}| = |\alpha|, \quad \arg \bar{\alpha} = -\arg \alpha. \tag{14}$$

Сумма и произведение сопряженных комплексных чисел являются действительными числами. В самом деле,

$$\left.\begin{aligned} \alpha + \bar{\alpha} &= 2a, \\ \alpha\bar{\alpha} &= a^2 + b^2 = |\alpha|^2. \end{aligned}\right\} \tag{15}$$

Последнее равенство показывает, что число $\alpha\bar{\alpha}$ даже положительно при $\alpha \neq 0$. В § 24 будет получена теорема, показывающая, что доказанное сейчас свойство сопряженных чисел является для них характерным.

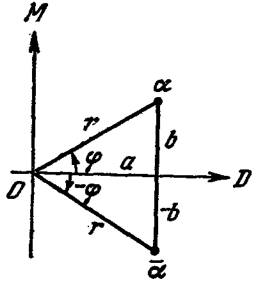


Рис. 7.

Равенство

$$(a - bi) + (c - di) = (a + c) - (b + d)\, i$$

показывает, что *число, сопряженное с суммой двух чисел, равно сумме чисел, сопряженных со слагаемыми:*

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}. \tag{16}$$

Аналогично, из равенства

$$(a - bi)(c - di) = (ac - bd) - (ad + bc)\, i$$

вытекает, что *число, сопряженное с произведением, равно произведению чисел, сопряженных с сомножителями:*

$$\overline{\alpha\beta} = \bar{\alpha} \cdot \bar{\beta}. \tag{17}$$

Непосредственная проверка показывает также справедливость формул

$$\overline{\alpha - \beta} = \bar{\alpha} - \bar{\beta}, \tag{18}$$

$$\overline{\left(\frac{\alpha}{\beta}\right)} = \frac{\bar{\alpha}}{\bar{\beta}}. \tag{19}$$

Докажем следующее утверждение: *если число α некоторым образом выражено через комплексные числа $\beta_1, \beta_2, \ldots, \beta_n$ при помощи сложения, умножения, вычитания и деления, то, заменяя в этом выражении все числа β_k их сопряженными, мы получим число, сопряженное с α*; в частности, если число α действительное, то оно не меняется при замене всех комплексных чисел β_k их сопряженными.

Будем доказывать это утверждение индукцией по n, так как при $n = 2$ оно вытекает из формул (16) — (19).

Пусть число α выражено через числа $\beta_1, \beta_2, \ldots, \beta_n$, не обязательно различные. В этом выражении указан определенный порядок,

в котором применяются операции сложения, умножения, вычитания и деления. Последним шагом будет применение одной из этих операций к числу γ_1, выраженному через числа $\beta_1, \beta_2, \ldots, \beta_k$, где $1 \leqslant k \leqslant n-1$, и к числу γ_2, выраженному через числа $\beta_{k+1}, \ldots, \beta_n$. По индуктивному предположению замена чисел $\beta_1, \beta_2, \ldots, \beta_k$ на сопряженные влечет за собой замену числа γ_1 на $\overline{\gamma_1}$, а замена чисел $\beta_{k+1}, \beta_{k+2}, \ldots, \beta_n$ на сопряженные заменяет γ_2 на $\overline{\gamma_2}$. Однако по одной из формул (16)—(19) переход от γ_1 и γ_2 к $\overline{\gamma_1}$ и $\overline{\gamma_2}$ превращает число α в $\overline{\alpha}$.

§ 19. Извлечение корня из комплексных чисел

Переходим к вопросу о возведении комплексных чисел в степень и извлечении из них корня. Для возведения числа $\alpha = a + bi$ в целую положительную степень n достаточно применить к выражению $(a+bi)^n$ формулу бинома Ньютона (эта формула справедлива и для комплексных чисел, так как ее доказательство основано лишь на законе дистрибутивности), а затем воспользоваться равенствами $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, откуда вообще

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i.$$

Если число α задано в тригонометрической форме, то при целом положительном n из формулы (4) предшествующего параграфа вытекает следующая формула, называемая *формулой Муавра*:

$$[r(\cos\varphi + i\sin\varphi)]^n = r^n(\cos n\varphi + i\sin n\varphi), \tag{1}$$

т. е. *при возведении комплексного числа в степень модуль возводится в эту степень, а аргумент умножается на показатель степени*. Формула (1) верна и для целых отрицательных показателей. Действительно, ввиду $\alpha^{-n} = (\alpha^{-1})^n$, достаточно применить формулу Муавра к числу α^{-1}, тригонометрическую форму которого дает формула (10) предшествующего параграфа.

Примеры.

1) $i^{37} = i, \quad i^{122} = -1;$

2) $(2+5i)^3 = 2^3 + 3\cdot 2^2\cdot 5i + 3\cdot 2\cdot 5^2 i^2 + 5^3 i^3 =$
$= 8 + 60i - 150 - 125i = -142 - 65i;$

3) $\left[\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)\right]^4 = (\sqrt{2})^4(\cos\pi + i\sin\pi) = -4;$

4) $\left[3\left(\cos\frac{\pi}{5} + i\sin\frac{\pi}{5}\right)\right]^{-3} =$
$= 3^{-3}\left[\cos\left(-\frac{3}{5}\pi\right) + i\sin\left(-\frac{3}{5}\pi\right)\right] = \frac{1}{27}\left(\cos\frac{7}{5}\pi + i\sin\frac{7}{5}\pi\right).$

Частный случай формулы Муавра, а именно равенство

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi,$$

позволяет легко получить формулы для синуса и косинуса кратного угла. Действительно, раскрывая левую часть этого равенства по формуле бинома и приравнивая отдельно действительные и мнимые части обеих частей равенства, мы получим:

$$\cos n\varphi = \cos^n\varphi - \binom{n}{2}\cos^{n-2}\varphi\cdot\sin^2\varphi + \binom{n}{4}\cos^{n-4}\varphi\cdot\sin^4\varphi - \ldots,$$

$$\sin n\varphi = \binom{n}{1}\cos^{n-1}\varphi\cdot\sin\varphi - \binom{n}{3}\cos^{n-3}\varphi\cdot\sin^3\varphi + \binom{n}{5}\cos^{n-5}\cdot\sin^5\varphi - \ldots;$$

здесь $\binom{n}{k}$ есть обычное обозначение биномиального коэффициента:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{1\cdot 2\cdot 3\ldots k}.$$

При $n=2$ мы приходим к известным формулам

$$\cos 2\varphi = \cos^2\varphi - \sin^2\varphi,$$

$$\sin 2\varphi = 2\cos\varphi\sin\varphi,$$

а при $n=3$ — к формулам

$$\cos 3\varphi = \cos^3\varphi - 3\cos\varphi\sin^2\varphi,$$

$$\sin 3\varphi = 3\cos^2\varphi\sin\varphi - \sin^3\varphi.$$

Извлечение корня из комплексных чисел представляет уже много больше трудностей. Начнем с извлечения квадратного корня из числа $\alpha = a+bi$. Мы не знаем пока, существует ли такое комплексное число, квадрат которого равен α. Предположим, что такое число $u+vi$ существует, т. е., употребляя обычную символику, можно написать

$$\sqrt{a+bi} = u+vi.$$

Из равенства

$$(u+vi)^2 = a+bi$$

следует

$$\left.\begin{aligned} u^2 - v^2 &= a,\\ 2uv &= b. \end{aligned}\right\} \qquad (2)$$

Возводя в квадрат обе части каждого из равенств (2), а затем складывая их, получаем:

$$(u^2-v^2)^2 + 4u^2v^2 = (u^2+v^2)^2 = a^2+b^2,$$

откуда

$$u^2+v^2 = +\sqrt{a^2+b^2};$$

положительный знак взят потому, что числа u и v действительные, и поэтому левая часть равенства положительная. Из этого равенства и из первого из равенств (2) получаем:

$$u^2=\frac{1}{2}\left(a+\sqrt{a^2+b^2}\right),$$

$$v^2=\frac{1}{2}\left(-a+\sqrt{a^2+b^2}\right).$$

Мы приходим, извлекая квадратные корни, к двум значениям для u, отличающимся друг от друга знаком, а также к двум значениям для v. Все эти значения будут действительными, так как квадратные корни будут извлекаться при любых a и b из положительных чисел. Полученные значения для u и v нельзя комбинировать между собой произвольным образом, так как, ввиду второго из равенств (2), знак произведения uv должен совпадать со знаком b. Это дает две возможные комбинации значений u и v, т. е. два числа вида $u+vi$, которые могут служить значениями квадратного корня из числа α; эти числа отличаются друг от друга знаком. Элементарная, хотя и громоздкая, проверка (возведением полученных чисел в квадрат, отдельно для случая $b>0$ и для случая $b<0$) показывает, что найденные нами числа действительно являются значениями квадратного корня из числа α. Таким образом, *извлечение квадратного корня из комплексного числа всегда возможно и дает два значения, отличающиеся друг от друга знаком.*

В частности, теперь делается возможным извлечение квадратного корня и из отрицательного действительного числа, причем значения этого корня будут чисто мнимыми. В самом деле, если $a<0$ и $b=0$, то $\sqrt{a^2+b^2}=-a$, так как этот корень должен быть положительным, а тогда $u^2=\frac{1}{2}(a-a)=0$, т. е. $u=0$, откуда $\sqrt{a}=\pm vi$.

Пример. Пусть $\alpha=21-20i$. Тогда $\sqrt{a^2+b^2}=\sqrt{441+400}=29$. Поэтому $u^2=\frac{1}{2}(21+29)=25$, $v^2=\frac{1}{2}(-21+29)=4$, откуда $u=\pm 5$, $v=\pm 2$. Знаки u и v должны быть различными ввиду отрицательности b, поэтому

$$\sqrt{21-20i}=\pm(5-2i).$$

Попытки извлечения из комплексных чисел, заданных в виде $a+bi$, корней более высокой степени, чем вторая, встречаются с непреодолимыми затруднениями. Так, если бы мы захотели таким же методом, как выше, извлечь из числа $a+bi$ кубичный корень, то должны были бы решить некоторое вспомогательное кубичное уравнение, чего мы пока не умеем и что в свою очередь требует, как мы узнаем в § 38, извлечения кубичного корня из комплексного числа. С другой стороны, тригонометрическая форма весьма хорошо

приспособлена для извлечения корней любой степени и, пользуясь ею, мы сейчас полностью исчерпаем этот вопрос.

Пусть нужно извлечь корень n-й степени из числа $\alpha = r(\cos\varphi + i\sin\varphi)$. Предположим, что это сделать можно и что в результате получается число $\rho(\cos\theta + i\sin\theta)$, т. е.

$$[\rho(\cos\theta + i\sin\theta)]^n = r(\cos\varphi + i\sin\varphi). \tag{3}$$

Тогда, по формуле Муавра, $\rho^n = r$, т. е. $\rho = \sqrt[n]{r}$, где в правой части стоит однозначно определенное положительное значение корня n-й степени из положительного действительного числа r. С другой стороны, аргумент левой части равенства (3) есть $n\theta$. Нельзя утверждать, однако, что $n\theta$ равно φ, так как эти углы могут в действительности отличаться на слагаемое, являющееся некоторым целым кратным числа 2π. Поэтому $n\theta = \varphi + 2k\pi$, где k — целое число, откуда

$$\theta = \frac{\varphi + 2k\pi}{n}.$$

Обратно, если мы берем число $\sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right)$, то при любом целом k, положительном или отрицательном, n-я степень этого числа равна α. Таким образом,

$$\sqrt[n]{r(\cos\varphi + i\sin\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right). \tag{4}$$

Давая k различные значения, мы не всегда будем получать различные значения искомого корня. Действительно, при

$$k = 0,\ 1,\ 2,\ \ldots,\ n-1 \tag{5}$$

мы получим n значений корня, которые все будут различными, так как увеличение k на единицу влечет за собой увеличение аргумента на $\frac{2\pi}{n}$. Пусть теперь k произвольно. Если $k = nq + r$, $0 \leqslant r \leqslant n-1$, то

$$\frac{\varphi+2k\pi}{n} = \frac{\varphi + 2(nq+r)\pi}{n} = \frac{\varphi+2r\pi}{n} + 2q\pi,$$

т. е. значение аргумента при нашем k отличается от значения аргумента при $k = r$ на число, кратное 2π; мы получаем, следовательно, такое же значение корня, как при значении k, равном r, т. е. входящем в систему (5).

Таким образом, *извлечение корня n-й степени из комплексного числа α всегда возможно и дает n различных значений. Все значения корня n-й степени расположены на окружности радиуса $\sqrt[n]{|\alpha|}$ с центром в нуле и делят эту окружность на n равных частей.*

В частности, корень n-й степени из действительного числа a также имеет n различных значений; действительных среди этих значений будет два, одно или ни одного в зависимости от знака a и четности n.

Примеры.

1) $$\beta=\sqrt[3]{2\left(\cos\frac{3}{4}\pi+i\sin\frac{3}{4}\pi\right)}=\sqrt[3]{2}\left(\cos\frac{\frac{3}{4}\pi+2k\pi}{3}+i\sin\frac{\frac{3}{4}\pi+2k\pi}{3}\right);$$

$$k=0:\quad \beta_0=\sqrt[3]{2}\left(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}\right);$$

$$k=1:\quad \beta_1=\sqrt[3]{2}\left(\cos\frac{11}{12}\pi+i\sin\frac{11}{12}\pi\right);$$

$$k=2:\quad \beta_2=\sqrt[3]{2}\left(\cos\frac{19}{12}\pi+i\sin\frac{19}{12}\pi\right).$$

2) $$\beta=\sqrt{i}=\sqrt{\cos\frac{\pi}{2}+i\sin\frac{\pi}{2}}=\cos\frac{\frac{\pi}{2}+2k\pi}{2}+i\sin\frac{\frac{\pi}{2}+2k\pi}{2};$$

$$\beta_0=\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}=\frac{\sqrt{2}}{2}+i\,\frac{\sqrt{2}}{2};\quad \beta_1=\cos\frac{5}{4}\pi+i\sin\frac{5}{4}\pi=-\beta_0.$$

3) $$\beta=\sqrt[3]{-8}=\sqrt[3]{8(\cos\pi+i\sin\pi)}=2\left(\cos\frac{\pi+2k\pi}{3}+i\sin\frac{\pi+2k\pi}{3}\right);$$

$$\beta_0=2\left(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}\right)=1+i\sqrt{3};$$

$$\beta_1=2(\cos\pi+i\sin\pi)=-2;$$

$$\beta_2=2\left(\cos\frac{5\pi}{3}+i\sin\frac{5\pi}{3}\right)=1-i\sqrt{3}.$$

Корни из единицы. Особенно важен случай извлечения корня n-й степени из числа 1. Этот корень имеет n значений, причем, ввиду равенства $1=\cos 0+i\sin 0$ и формулы (4), все эти значения или, как мы будем говорить, все *корни n-й степени из единицы*, даются формулой

$$\sqrt[n]{1}=\cos\frac{2k\pi}{n}+i\sin\frac{2k\pi}{n};\quad k=0,\ 1,\ \ldots,\ n-1. \tag{6}$$

Действительные значения корня n-й степени из единицы получаются из формулы (6) при значениях $k=0$ и $\frac{n}{2}$, если n четно, и при $k=0$, если n нечетно. На комплексной плоскости корни n-й степени из единицы расположены на окружности единичного круга и делят ее на n равных дуг; одной из точек деления служит число 1. Отсюда следует, что те из корней n-й степени из единицы, которые не являются действительными, расположены симметрично относительно действительной оси, т. е. попарно сопряжены.

Квадратный корень из единицы имеет два значения: 1 и -1, корень четвертой степени из единицы — четыре значения: 1, -1, i и $-i$. Для дальнейшего полезно запомнить значения *кубичного корня из единицы*. Это будут, ввиду (6), числа $\cos\frac{2k\pi}{3}+i\sin\frac{2k\pi}{3}$, где $k=0$, 1, 2, т. е., кроме самой единицы, также сопряженные между собою числа

$$\left.\begin{aligned} \varepsilon_1 &= \cos\frac{2\pi}{3}+i\sin\frac{2\pi}{3} = -\frac{1}{2}+i\frac{\sqrt{3}}{2},\\ \varepsilon_2 &= \cos\frac{4\pi}{3}+i\sin\frac{4\pi}{3} = -\frac{1}{2}-i\frac{\sqrt{3}}{2}. \end{aligned}\right\} \tag{7}$$

Все значения корня n-й степени из комплексного числа α можно получить умножением одного из этих значений на все корни n-й степени из единицы. Действительно, пусть β будет одно из значений корня n-й степени из числа α, т. е. $\beta^n=\alpha$, а ε — произвольное значение корня n-й степени из единицы, т. е. $\varepsilon^n=1$. Тогда $(\beta\varepsilon)^n=\beta^n\varepsilon^n=\alpha$, т. е. $\beta\varepsilon$ также будет одним из значений для $\sqrt[n]{\alpha}$. Умножая β на каждый из корней n-й степени из единицы, мы получаем n различных значений корня n-й степени из числа α, т. е. все значения этого корня.

Примеры. 1) Одно из значений кубичного корня из -8 есть -2. Два других будут, ввиду (7), числа $-2\varepsilon_1=1-i\sqrt{3}$ и $-2\varepsilon_2=1+i\sqrt{3}$ (см. выше пример 3)).

2) $\sqrt[4]{81}$ имеет четыре значения: 3, -3, $3i$, $-3i$.

Произведение двух корней n-й степени из единицы само есть корень n-й степени из единицы. Действительно, если $\varepsilon^n=1$ и $\eta^n=1$, то $(\varepsilon\eta)^n=\varepsilon^n\eta^n=1$. Далее, *число, обратное корню n-й степени из единицы, само есть такой же корень.* В самом деле, пусть $\varepsilon^n=1$. Тогда из $\varepsilon\cdot\varepsilon^{-1}=1$ следует $\varepsilon^n\cdot(\varepsilon^{-1})^n=1$, т. е. $(\varepsilon^{-1})^n=1$. Вообще, *всякая степень корня n-й степени из единицы есть также корень n-й степени из единицы.*

Всякий корень k-й степени из единицы будет также корнем l-й степени из единицы для всякого l, кратного k. Отсюда следует, что если мы будем рассматривать всю совокупность корней n-й степени из единицы, то некоторые из этих корней уже будут корнями n'-й степени из единицы для некоторых n', являющихся делителями числа n. Для всякого n существуют, однако, такие корни n-й степени из единицы, которые не являются корнями из единицы никакой меньшей степени. Такие корни называются *первообразными корнями n-й степени из единицы.* Их существование вытекает из формулы (6): если значение корня, соответствующее данному значению k, мы обозначим через ε_k (так что $\varepsilon_0=1$), то на основании формулы Муавра (1)

$$\varepsilon_1^k=\varepsilon_k.$$

Никакая степень числа ε_1, меньшая, чем n-я, не будет, следовательно, равна 1, т. е. $\varepsilon_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$ является первообразным корнем.

Корень n-й степени из единицы ε тогда и только тогда будет первообразным, если его степени ε^k, $k = 0, 1, \ldots, n-1$, различны, т. е. если ими исчерпываются все корни n-й степени из единицы.

Действительно, если все указанные степени числа ε различны, то ε будет, очевидно, первообразным корнем n-й степени. Если же, например, $\varepsilon^k = \varepsilon^l$ при $0 \leqslant k < l \leqslant n-1$, то $\varepsilon^{l-k} = 1$, т. е., ввиду неравенств $1 \leqslant l-k \leqslant n-1$, корень ε не будет первообразным.

Число ε_1, найденное выше, в общем случае — не единственный первообразный корень n-й степени. Для разыскания всех этих корней служит следующая теорема.

Если ε есть первообразный корень n-й степени из единицы, то число ε^k тогда и только тогда будет первообразным корнем n-й степени, если k взаимно просто с n.

В самом деле, пусть d будет наибольшим общим делителем чисел k и n. Если $d > 1$ и $k = dk'$, $n = dn'$, то

$$(\varepsilon^k)^{n'} = \varepsilon^{kn'} = \varepsilon^{k'n} = (\varepsilon^n)^{k'} = 1,$$

т. е. корень ε^k оказался корнем n'-й степени из единицы.

Пусть, с другой стороны, $d = 1$ и пусть, вместе с тем, число ε^k оказывается корнем m-й степени из единицы, $1 \leqslant m < n$. Таким образом

$$(\varepsilon^k)^m = \varepsilon^{km} = 1.$$

Так как число ε — первообразный корень n-й степени из единицы, т. е. лишь его степени с показателями, кратными n, могут быть равными единице, то число km будет кратным n. Отсюда вытекает, однако, так как $1 \leqslant m < n$, что числа k и n не могут быть взаимно простыми в противоречие с предположением.

Таким образом, число первообразных корней n-й степени из единицы равно числу целых положительных чисел k, меньших n и взаимно простых с ним. Выражение для этого числа, обычно обозначаемого через $\varphi(n)$, можно найти в любом курсе теории чисел.

Если p — простое число, то первообразными корнями p-й степени из единицы будут все эти корни, кроме самой единицы. С другой стороны, среди корней четвертой степени из единицы первообразными будут i и $-i$, но не 1 и -1.

ГЛАВА ПЯТАЯ

МНОГОЧЛЕНЫ И ИХ КОРНИ

§ 20. Операции над многочленами

Содержание первых двух глав книги, а именно — теория определителей и теория систем линейных уравнений, возникло в качестве непосредственного развития того направления в школьном курсе алгебры, которое, начинаясь от одного уравнения первой степени с одним неизвестным, вело к системам двух и трех уравнений первой степени с двумя и, соответственно, тремя неизвестными. Другое направление в элементарной алгебре, воспринимавшееся там как еще более значительное, состояло в переходе от уравнения первой степени с одним неизвестным к произвольному квадратному уравнению снова с одним неизвестным, а затем и к некоторым частным типам уравнений третьей и четвертой степени. Это направление вырастает в весьма большой и содержательный раздел высшей алгебры, посвященный изучению произвольных уравнений любой n-й степени с одним неизвестным. К этому разделу алгебры, исторически самому раннему, относятся как настоящая глава, так и некоторые из дальнейших глав книги.

Общий вид уравнения n-й степени (где n — некоторое целое положительное число) есть

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0. \qquad (1)$$

Коэффициенты $a_0, a_1, \ldots, a_{n-1}, a_n$ этого уравнения мы будем считать произвольными комплексными числами, причем *старший коэффициент* a_0 должен быть отличным от нуля.

Если написано уравнение (1), то всегда предполагается, что требуется его решить. Иными словами, требуется найти такие числовые значения для неизвестного x, которые удовлетворяют этому уравнению, т. е. после подстановки вместо неизвестного и выполнения всех указанных операций обращают левую часть уравнения (1) в нуль.

Целесообразно, однако, заменить задачу решения уравнения (1) более общей задачей изучения левой части этого уравнения

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \qquad (2)$$

называемой *многочленом* (или *полиномом*) *n-й степени от неизвестного* x. Мы выбираем первый из этих терминов; следует твердо помнить, что теперь многочленом называется лишь выражение вида (2), т. е. лишь сумма целых неотрицательных степеней неизвестного x, взятых с некоторыми числовыми коэффициентами, а не любая сумма одночленов, как это было в элементарной алгебре. В частности, мы не будем считать многочленами такие выражения, которые содержат неизвестное x с отрицательными или дробными показателями, например, $2x^2-\frac{1}{x}+3$, или $ax^{-3}+bx^{-2}+cx^{-1}+d+ex+fx^2$, или же $x^{\frac{1}{2}}+1$. Для сокращенной записи многочленов будут употребляться символы $f(x)$, $g(x)$, $\varphi(x)$ и т. д.

Два многочлена $f(x)$ и $g(x)$ будут считаться *равными* (или *тождественно равными*), $f(x)=g(x)$, в том случае, если равны их коэффициенты при одинаковых степенях неизвестного. В частности, никакой многочлен, хотя бы один коэффициент которого отличен от нуля, не может быть равным нулю, и поэтому знак равенства, употребляемый в записи уравнения n-й степени (1), не имеет никакого отношения к определенному сейчас равенству многочленов. Знак $=$, связывающий многочлены, следует в дальнейшем всегда понимать в смысле тождественного равенства этих многочленов.

Таким образом, на многочлен n-й степени (2) следует смотреть как на некоторое формальное выражение, вполне определяемое набором своих коэффициентов $a_0, a_1, \ldots, a_n$, где $a_0 \neq 0$. Точный смысл этих слов будет выяснен много позже, в гл. 10. Заметим, что, помимо записи многочлена в виде (2), т. е. по убывающим степеням неизвестного x, будут допускаться и другие его записи, получающиеся из (2) перестановкой слагаемых, например, запись, по возрастающим степеням неизвестного.

Конечно, на многочлен (2) можно было бы смотреть и с точки зрения математического анализа, т. е. считать его комплексной функцией комплексного переменного x. Следует учесть, однако, что две функции считаются равными в том случае, если равны их значения при всех значениях переменного x. Ясно, что два многочлена, равные в указанном выше формально-алгебраическом смысле, будут равны и как функции от x. Обратное будет доказано, однако, лишь в § 24. После этого алгебраическая и теоретико-функциональная точки зрения на понятие многочлена с числовыми коэффициентами на самом деле станут равносильными, пока же мы должны каждый раз указывать, какой именно смысл придается понятию многочлена. В настоящем и двух следующих параграфах мы будем смотреть на многочлен как на формально-алгебраическое выражение.

Существуют, понятно, многочлены n-й степени для любого натурального числа n. Рассматривая всевозможные такие многочлены,

мы, помимо многочленов первой степени, квадратных, кубичных и т. д., встретимся и с *многочленами нулевой степени*, т. е. с отличными от нуля комплексными числами. Число нуль также будет считаться многочленом; это будет единственный многочлен, степень которого не определена.

Сейчас мы определим для многочленов с комплексными коэффициентами операции сложения и умножения. Эти операции будут введены по образцу операций над многочленами с действительными коэффициентами, известных читателю из курса элементарной алгебры.

Если даны многочлены $f(x)$ и $g(x)$ с комплексными коэффициентами, записанные для удобства по возрастающим степеням x:

$$f(x)=a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n,\quad a_n\neq 0,$$
$$g(x)=b_0+b_1x+\ldots+b_{s-1}x^{s-1}+b_sx^s,\quad b_s\neq 0,$$

и если, например, $n\geqslant s$, то их *суммой* называется многочлен

$$f(x)+g(x)=c_0+c_1x+\ldots+c_{n-1}x^{n-1}+c_nx^n,$$

коэффициенты которого получаются сложением коэффициентов многочленов $f(x)$ и $g(x)$, стоящих при одинаковых степенях неизвестного, т. е.

$$c_i=a_i+b_i,\quad i=0,\ 1,\ \ldots,\ n,\tag{3}$$

причем при $n>s$ коэффициенты $b_{s+1},\ b_{s+2},\ \ldots,\ b_n$ следует считать равными нулю. Степень суммы будет равна n, если n больше s, но при $n=s$ она может случайно оказаться меньше n, а именно в случае $b_n=-a_n$.

Произведением многочленов $f(x)$ и $g(x)$ называется многочлен

$$f(x)\cdot g(x)=d_0+d_1x+\ldots+d_{n+s-1}x^{n+s-1}+d_{n+s}x^{n+s},$$

коэффициенты которого определяются следующим образом:

$$d_i=\sum_{k+l=i}a_kb_l,\quad i=0,\ 1,\ \ldots,\ n+s-1,\ n+s,\tag{4}$$

т. е. коэффициент d_i есть результат перемножения таких коэффициентов многочленов $f(x)$ и $g(x)$, сумма индексов которых равна i, и сложения всех таких произведений; в частности, $d_0=a_0b_0$, $d_1=a_0b_1+a_1b_0$, $\ldots$, $d_{n+s}=a_nb_s$. Из последнего равенства вытекает неравенство $d_{n+s}\neq 0$ и поэтому *степень произведения двух многочленов равна сумме степеней этих многочленов.*

Отсюда следует, что *произведение многочленов, отличных от нуля, никогда не будет равным нулю.*

Какими свойствами обладают введенные нами операции для многочленов? Коммутативность и ассоциативность сложения немедленно вытекают из справедливости этих свойств для сложения чисел, так как складываются коэффициенты при каждой степени неизвестного отдельно. Вычитание оказывается выполни-

мым: роль нуля играет число нуль, включенное нами в число многочленов, а противоположным для записанного выше многочлена $f(x)$ будет многочлен

$$-f(x) = -a_0 - a_1 x - \ldots - a_{n-1}x^{n-1} - a_n x^n.$$

Коммутативность умножения вытекает из коммутативности умножения чисел и того факта, что в определении произведения многочленов коэффициенты обоих множителей $f(x)$ и $g(x)$ используются совершенно равноправным образом. Ассоциативность умножения доказывается следующим образом: если, помимо записанных выше многочленов $f(x)$ и $g(x)$, дан еще многочлен

$$h(x) = c_0 + c_1 x + \ldots + c_{t-1}x^{t-1} + c_t x^t, \quad c_t \neq 0,$$

то коэффициентом при x^i, $i = 0, 1, \ldots, n+s+t$ в произведении $[f(x)\,g(x)]\,h(x)$ будет служить число

$$\sum_{j+m=i}\left(\sum_{k+l=j} a_k b_l\right) c_m = \sum_{k+l+m=i} a_k b_l c_m,$$

а в произведении $f(x)\,[g(x)\,h(x)]$ — равное ему число

$$\sum_{k+j=i} a_k \left(\sum_{l+m=j} b_l c_m\right) = \sum_{k+l+m=i} a_k b_l c_m.$$

Наконец, справедливость закона дистрибутивности вытекает из равенства

$$\sum_{k+l=i} (a_k + b_k)\,c_l = \sum_{k+l=i} a_k c_l + \sum_{k+l=i} b_k c_l,$$

так как левая часть этого равенства является коэффициентом при x^i в многочлене $[f(x) + g(x)]\,h(x)$, а правая часть — коэффициентом при той же степени неизвестного в многочлене $f(x)\,h(x) + g(x)\,h(x)$.

Заметим, что роль единицы при умножении многочленов играет число 1, рассматриваемое как многочлен нулевой степени. С другой стороны, *многочлен $f(x)$ тогда и только тогда обладает обратным многочленом $f^{-1}(x)$,*

$$f(x)\,f^{-1}(x) = 1, \tag{5}$$

если $f(x)$ является многочленом нулевой степени. Действительно, если $f(x)$ является отличным от нуля числом a, то обратным многочленом служит для него число a^{-1}. Если же $f(x)$ имеет степень $n \geqslant 1$, то степень левой части равенства (5), если бы многочлен $f^{-1}(x)$ существовал, была бы не меньше n, в то время как справа стоит многочлен нулевой степени.

Отсюда вытекает, что *для умножения многочленов обратная операция — деление — не существует.* В этом отношении система всех многочленов с комплексными коэффициентами напоминает систему всех целых чисел. Эта аналогия проявляется и в том, что для

многочленов, как и для целых чисел, существует алгоритм деления с остатком. Этот алгоритм для случая многочленов с действительными коэффициентами известен читателю еще из элементарной алгебры. Так как, однако, мы рассматриваем теперь случай многочленов с комплексными коэффициентами, следует еще раз дать все относящиеся сюда формулировки и привести доказательства.

Для любых двух многочленов $f(x)$ *и* $g(x)$ *можно найти такие многочлены* $q(x)$ *и* $r(x)$, *что*

$$f(x) = g(x)\,q(x) + r(x), \tag{6}$$

причем степень $r(x)$ *меньше степени* $g(x)$ *или же* $r(x) = 0$. *Многочлены* $q(x)$ *и* $r(x)$, *удовлетворяющие этому условию, определяются однозначно.*

Докажем сперва вторую половину теоремы. Пусть существуют еще многочлены $\overline{q}(x)$ и $\overline{r}(x)$, также удовлетворяющие равенству

$$f(x) = g(x)\,\overline{q}(x) + \overline{r}(x), \tag{7}$$

причем степень $\overline{r}(x)$ снова меньше степени $g(x)$[1]. Приравнивая друг другу правые части равенств (6) и (7), получим:

$$g(x)\,[q(x) - \overline{q}(x)] = \overline{r}(x) - r(x).$$

Степень правой части этого равенства меньше степени $g(x)$, степень же левой части была бы при $q(x) - \overline{q}(x) \neq 0$ больше или равна степени $g(x)$. Поэтому должно быть $q(x) - \overline{q}(x) = 0$, т. е. $q(x) = \overline{q}(x)$, а тогда и $r(x) = \overline{r}(x)$, что и требовалось доказать.

Переходим к доказательству первой половины теоремы. Пусть многочлены $f(x)$ и $g(x)$ имеют соответственно степени n и s. Если $n < s$, то можно положить $q(x) = 0$, $r(x) = f(x)$. Если же $n \geqslant s$, то воспользуемся тем же методом, каким в элементарной алгебре производилось деление многочленов с действительными коэффициентами, расположенных по убывающим степеням неизвестного. Пусть

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \qquad a_0 \neq 0,$$
$$g(x) = b_0x^s + b_1x^{s-1} + \ldots + b_{s-1}x + b_s, \qquad b_0 \neq 0.$$

Полагая

$$f(x) - \frac{a_0}{b_0}x^{n-s}g(x) = f_1(x), \tag{8}$$

мы получим многочлен, степень которого меньше n. Обозначим эту степень через n_1, а старший коэффициент многочлена $f_1(x)$ — через a_{10}. Положим, далее, если все еще $n_1 \geqslant s$,

$$f_1(x) - \frac{a_{10}}{b_0}x^{n_1-s}g(x) = f_2(x), \tag{8$_1$}$$

[1]) Или же $\overline{r}(x) = 0$. В дальнейшем этот случай не будет оговариваться.

обозначим через n_2 — степень, а через a_{20} — старший коэффициент многочлена $f_2(x)$, положим затем

$$f_2(x)-\frac{a_{20}}{b_0}x^{n_2-s}g(x)=f_3(x), \qquad (8_2)$$

и т. д.

Так как степени многочленов $f_1(x)$, $f_2(x)$, ... убывают, $n>n_1>$ $>n_2>\ldots$, то мы дойдем после конечного числа шагов до такого многочлена $f_k(x)$,

$$f_{k-1}(x)-\frac{a_{k-1,0}}{b_0}x^{n_{k-1}-s}g(x)=f_k(x), \qquad (8_{k-1})$$

степень которого n_k меньше s, после чего наш процесс останавливается. Складывая теперь равенства (8), (8_1), ..., (8_{k-1}), мы получим:

$$f(x)-\left(\frac{a_0}{b_0}x^{n-s}+\frac{a_{10}}{b_0}x^{n_1-s}+\ldots+\frac{a_{k-1,0}}{b_0}x^{n_{k-1}-s}\right)g(x)=f_k(x),$$

т. е. многочлены

$$q(x)=\frac{a_0}{b_0}x^{n-s}+\frac{a_{10}}{b_0}x^{n_1-s}+\ldots+\frac{a_{k-1,0}}{b_0}x^{n_{k-1}-s},$$
$$r(x)=f_k(x)$$

действительно удовлетворяют равенству (6), причем степень $r(x)$ на самом деле меньше степени $g(x)$.

Заметим, что многочлен $q(x)$ называется *частным* от деления $f(x)$ на $g(x)$, а $r(x)$ — *остатком* от этого деления.

Из рассмотрения алгоритма деления с остатком легко устанавливается, что если $f(x)$ и $g(x)$ *являются многочленами с действительными коэффициентами, то коэффициенты всех многочленов* $f_1(x)$, $f_2(x)$, ..., *а поэтому и коэффициенты частного* $q(x)$ *и остатка* $r(x)$ *будут действительными.*

§ 21. Делители. Наибольший общий делитель

Пусть даны ненулевые многочлены $f(x)$ и $\varphi(x)$ с комплексными коэффициентами. Если остаток от деления $f(x)$ на $\varphi(x)$ равен нулю, т. е., как говорят, $f(x)$ *делится* (или *нацело делится*) на $\varphi(x)$, то многочлен $\varphi(x)$ называется *делителем* многочлена $f(x)$.

Многочлен $\varphi(x)$ *тогда и только тогда будет делителем многочлена* $f(x)$, *если существует многочлен* $\psi(x)$, *удовлетворяющий равенству*

$$f(x)=\varphi(x)\psi(x). \qquad (1)$$

В самом деле, если $\varphi(x)$ является делителем для $f(x)$, то в качестве $\psi(x)$ следует взять частное от деления $f(x)$ на $\varphi(x)$. Обратно, пусть многочлен $\psi(x)$, удовлетворяющий равенству (1), существует.

Из доказанной в предшествующем параграфе единственности многочленов $q(x)$ и $r(x)$, удовлетворяющих равенству

$$f(x) = \varphi(x) q(x) + r(x)$$

и условию, что степень $r(x)$ меньше степени $\varphi(x)$, в нашем случае следует, что частное от деления $f(x)$ на $\varphi(x)$ равно $\psi(x)$, а остаток равен нулю.

Понятно, что если равенство (1) имеет место, то $\psi(x)$ также будет делителем для $f(x)$. Очевидно, далее, что степень $\varphi(x)$ не больше степени $f(x)$.

Заметим, что если многочлен $f(x)$ и его делитель $\varphi(x)$ имеют оба рациональные или действительные коэффициенты, то и многочлен $\psi(x)$ также будет иметь рациональные, или, соответственно, действительные коэффициенты, так как он разыскивается при помощи алгоритма деления. Конечно, многочлен с рациональными или действительными коэффициентами может обладать и такими делителями, не все коэффициенты которых рациональны или, соответственно, действительны. Это показывает, например, равенство

$$x^2 + 1 = (x - i)(x + i).$$

Укажем некоторые основные свойства делимости многочленов, которые найдут в дальнейшем многочисленные применения.

I. *Если $f(x)$ делится на $g(x)$, а $g(x)$ делится на $h(x)$, то $f(x)$ будет делиться на $h(x)$.*

В самом деле, по условию $f(x) = g(x)\varphi(x)$ и $g(x) = h(x)\psi(x)$, а поэтому $f(x) = h(x)[\psi(x)\varphi(x)]$.

II. *Если $f(x)$ и $g(x)$ делятся на $\varphi(x)$, то их сумма и разность также делятся на $\varphi(x)$.*

Действительно, из равенств $f(x) = \varphi(x)\psi(x)$ и $g(x) = \varphi(x)\chi(x)$ вытекает $f(x) \pm g(x) = \varphi(x)[\psi(x) \pm \chi(x)]$.

III. *Если $f(x)$ делится на $\varphi(x)$, то произведение $f(x)$ на любой многочлен $g(x)$ также будет делиться на $\varphi(x)$.*

Действительно, если $f(x) = \varphi(x)\psi(x)$, то $f(x)g(x) = \varphi(x)[\psi(x)g(x)]$.

Из II и III вытекает следующее свойство:

IV. *Если каждый из многочленов $f_1(x), f_2(x), \ldots, f_k(x)$ делится на $\varphi(x)$, то на $\varphi(x)$ будет делиться и многочлен*

$$f_1(x)g_1(x) + f_2(x)g_2(x) + \ldots + f_k(x)g_k(x),$$

где $g_1(x), g_2(x), \ldots, g_k(x)$ — произвольные многочлены.

V. *Всякий многочлен $f(x)$ делится на любой многочлен нулевой степени.*

Действительно, если $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$, а c — произвольное число, не равное нулю, т. е. произвольный многочлен нулевой степени, то

$$f(x) = c\left(\frac{a_0}{c}x^n + \frac{a_1}{c}x^{n-1} + \ldots + \frac{a_n}{c}\right).$$

VI. *Если* $f(x)$ *делится на* $\varphi(x)$, *то* $f(x)$ *делится и на* $c\varphi(x)$, *где* c — *произвольное число, отличное от нуля.*

В самом деле, из равенства $f(x)=\varphi(x)\psi(x)$ следует равенство $f(x)=[c\varphi(x)]\cdot[c^{-1}\psi(x)]$.

VII. *Многочлены* $cf(x)$, $c\neq 0$, *и только они будут делителями многочлена* $f(x)$, *имеющими такую же степень, что и* $f(x)$.

Действительно, $f(x)=c^{-1}[cf(x)]$, т. е. $f(x)$ делится на $cf(x)$. Если, с другой стороны, $f(x)$ делится на $\varphi(x)$, причем степени $f(x)$ и $\varphi(x)$ совпадают, то степень частного от деления $f(x)$ на $\varphi(x)$ должна быть равной нулю, т. е. $f(x)=d\varphi(x)$, $d\neq 0$, откуда $\varphi(x)=d^{-1}f(x)$.

Отсюда вытекает следующее свойство:

VIII. *Тогда и только тогда многочлены* $f(x)$, $g(x)$ *одновременно делятся друг на друга, если* $g(x)=cf(x)$, $c\neq 0$.

Наконец, из VIII и I вытекает свойство

IX. *Всякий делитель одного из двух многочленов* $f(x)$, $cf(x)$, *где* $c\neq 0$, *будет делителем и для другого многочлена.*

Наибольший общий делитель. Пусть даны произвольные многочлены $f(x)$ и $g(x)$. Многочлен $\varphi(x)$ будет называться *общим делителем* для $f(x)$ и $g(x)$, если он служит делителем для каждого из этих многочленов. Свойство V (см. выше) показывает, что к числу общих делителей многочленов $f(x)$ и $g(x)$ принадлежат все многочлены нулевой степени. Если других общих делителей эти два многочлена не имеют, то они называются *взаимно простыми.*

В общем же случае многочлены $f(x)$ и $g(x)$ могут обладать делителями, зависящими от x, и мы хотим ввести понятие о *наибольшем* общем делителе этих многочленов.

Было бы неудобным принять такое определение, по которому наибольший общий делитель многочленов $f(x)$ и $g(x)$ есть их общий делитель наибольшей степени. С одной стороны, мы не знаем пока, не будут ли $f(x)$ и $g(x)$ обладать многими различными общими делителями наибольшей степени, отличающимися друг от друга не только на множитель нулевой степени, т. е. не содержит ли это определение слишком большой неопределенности. С другой стороны, читатель уже встречался в элементарной арифметике с задачей разыскания наибольшего общего делителя целых чисел и знает, что наибольший общий делитель 6 целых чисел 12 и 18 не только является наибольшим среди общих делителей этих чисел, но даже делится на любой другой их общий делитель; действительно, другими общими делителями чисел 12 и 18 будут числа 1, 2, 3, —1, —2, —3, —6.

Мы примем поэтому для случая многочленов такое определение:

Наибольшим общим делителем отличных от нуля многочленов $f(x)$ и $g(x)$ называется такой многочлен $d(x)$, который является их общим делителем и, вместе с тем, сам делится на любой другой общий делитель этих многочленов. Обозначается наибольший общий делитель многочленов $f(x)$ и $g(x)$ символом $(f(x),\ g(x))$.

Это определение оставляет открытым вопрос, существует ли наибольший общий делитель для любых многочленов $f(x)$ и $g(x)$. Сейчас на этот вопрос будет дан положительный ответ. Одновременно будет дан метод для практического разыскания наибольшего общего делителя данных многочленов. Понятно, что мы не можем перенести сюда тот способ, каким обычно разыскивается наибольший общий делитель целых чисел, так как пока не имеем для многочленов ничего аналогичного разложению целого числа в произведение простых множителей. Для целых чисел существует, однако, и другой способ, называемый *алгоритмом последовательного деления* или *алгоритмом Евклида*; этот способ вполне применим и к многочленам.

Алгоритм Евклида для многочленов состоит в следующем. Пусть даны многочлены $f(x)$ и $g(x)$. Делим $f(x)$ на $g(x)$ и получаем, вообще говоря, некоторый остаток $r_1(x)$. Делим затем $g(x)$ на $r_1(x)$ и получаем остаток $r_2(x)$, делим $r_1(x)$ на $r_2(x)$ и т. д. Так как степени остатков все время понижаются, то в этой цепочке последовательных делений мы должны дойти до такого места, на котором деление совершится нацело и поэтому процесс остановится. *Тот остаток $r_k(x)$, на который нацело делится предыдущий остаток $r_{k-1}(x)$, и будет наибольшим общим делителем многочленов $f(x)$ и $g(x)$.*

Для доказательства запишем изложенное в предыдущем абзаце в виде следующей цепочки равенств:

$$\left.\begin{aligned} f(x) &= g(x)\,q_1(x) + r_1(x), \\ g(x) &= r_1(x)\,q_2(x) + r_2(x), \\ r_1(x) &= r_2(x)\,q_3(x) + r_3(x), \\ &\ldots\ldots\ldots\ldots \\ r_{k-3}(x) &= r_{k-2}(x)\,q_{k-1}(x) + r_{k-1}(x), \\ r_{k-2}(x) &= r_{k-1}(x)\,q_k(x) + r_k(x), \\ r_{k-1}(x) &= r_k(x)\,q_{k+1}(x). \end{aligned}\right\} \qquad (2)$$

Последнее равенство показывает, что $r_k(x)$ служит делителем для $r_{k-1}(x)$. Отсюда следует, что оба слагаемых правой части предпоследнего равенства делятся на $r_k(x)$, а поэтому $r_k(x)$ будет делителем и для $r_{k-2}(x)$. Далее, таким же путем, поднимаясь вверх, мы получим, что $r_k(x)$ является делителем и для $r_{k-3}(x), \ldots, r_2(x), r_1(x)$. Отсюда, ввиду второго равенства, будет следовать, что $r_k(x)$ служит делителем для $g(x)$, а поэтому, на основании первого равенства,— и для $f(x)$. Таким образом, $r_k(x)$ является общим делителем для $f(x)$ и $g(x)$.

Возьмем теперь произвольный общий делитель $\varphi(x)$ многочленов $f(x)$ и $g(x)$. Так как левая часть и первое слагаемое правой части первого из равенств (2) делятся на $\varphi(x)$, то $r_1(x)$ также будет делиться на $\varphi(x)$. Переходя ко второму и следующему равенствам,

мы таким же способом получим, что на $\varphi(x)$ делятся многочлены $r_2(x)$, $r_3(x)$, ... Наконец, если уже будет доказано, что $r_{k-2}(x)$ и $r_{k-1}(x)$ делятся на $\varphi(x)$, то из предпоследнего равенства мы получим, что $r_k(x)$ делится на $\varphi(x)$. Таким образом, $r_k(x)$ на самом деле будет наибольшим общим делителем для $f(x)$ и $g(x)$.

Мы доказали, следовательно, что любые два многочлена обладают наибольшим общим делителем, и получили способ для его вычисления. Этот способ показывает, что *если многочлены $f(x)$ и $g(x)$ имеют оба рациональные или действительные коэффициенты, то и коэффициенты их наибольшего общего делителя также будут рациональными или, соответственно, действительными,* хотя, конечно, у этих многочленов могут существовать и такие делители, не все коэффициенты которых рациональны (действительны). Так, многочлены с рациональными коэффициентами

$$f(x)=x^3-3x^2-2x+6,\quad g(x)=x^3+x^2-2x-2$$

имеют наибольшим общим делителем многочлен с рациональными коэффициентами x^2-2, хотя у них есть общий делитель $x-\sqrt{2}$, не все коэффициенты которого рациональны.

Если $d(x)$ есть наибольший общий делитель многочленов $f(x)$ и $g(x)$, то, как показывают свойства VIII и IX (см. выше), в качестве наибольшего общего делителя этих многочленов можно было бы выбрать также многочлен $cd(x)$, где c — произвольное число, отличное от нуля. Иными словами, *наибольший общий делитель двух многочленов определен лишь с точностью до множителя нулевой степени.* Ввиду этого можно условиться, что старший коэффициент наибольшего общего делителя двух многочленов будет всегда считаться равным единице. Используя это условие, можно сказать, что *два многочлена тогда и только тогда взаимно просты, если их наибольший общий делитель равен единице.* В самом деле, в качестве наибольшего общего делителя двух взаимно простых многочленов можно взять любое число, отличное от нуля, но, умножая его на обратный элемент, мы получим единицу.

Пример. Найти наибольший общий делитель многочленов

$$f(x)=x^4+3x^3-x^2-4x-3,\quad g(x)=3x^3+10x^2+2x-3.$$

Применяя алгоритм Евклида к многочленам с целыми коэффициентами, мы можем, чтобы избежать дробных коэффициентов, умножить делимое или сократить делитель на любое не равное нулю число, причем не только начиная какое-либо из последовательных делений, но и в процессе самого этого деления. Это будет приводить, понятно, к искажению частного, но интересующие нас остатки будут приобретать лишь некоторый множитель нулевой степени, что, как мы знаем, при разыскании наибольшего общего делителя допускается.

Делим $f(x)$ на $g(x)$, предварительно умножив $f(x)$ на 3:

$$\begin{array}{l|l} 3x^4+9x^3-3x^2-12x-9 & 3x^3+10x^2+2x-3 \\ \underline{3x^4+10x^3+2x^2-3x} & x+1 \\ \quad -x^3-5x^2-9x-9 & \end{array}$$

(умножаем на -3)

$$\begin{array}{r} 3x^3+15x^2+27x+27 \\ \underline{3x^3+10x^2+2x-3} \\ 5x^2+25x+30. \end{array}$$

Таким образом, первый остаток, после сокращения на 5, будет $r_1(x)=$ $=x^2+5x+6$. Делим на него многочлен $g(x)$:

$$\begin{array}{r|l} 3x^3+10x^2+2x-3 & x^2+5x+6 \\ \underline{3x^3+15x^2+18x\quad} & 3x-5 \\ -5x^2-16x-3 & \\ \underline{-5x^2-25x-30} & \\ 9x+27. & \end{array}$$

Вторым остатком, после сокращения на 9, будет, следовательно, $r_2(x)=x+3$. Так как

$$r_1(x)=r_2(x)(x+2),$$

то $r_2(x)$ будет тем последним остатком, на который нацело делится предшествующий остаток. Он будет, таким образом, искомым наибольшим общим делителем:

$$(f(x),\ g(x))=x+3.$$

Используем алгоритм Евклида для доказательства следующей теоремы:

Если $d(x)$ есть наибольший общий делитель многочленов $f(x)$ и $g(x)$, то можно найти такие многочлены $u(x)$ и $v(x)$, что

$$f(x)u(x)+g(x)v(x)=d(x). \qquad (3)$$

Можно считать при этом, если степени многочленов $f(x)$ и $g(x)$ больше нуля, что степень $u(x)$ меньше степени $g(x)$, а степень $v(x)$ меньше степени $f(x)$.

Доказательство основано на равенствах (2). Если мы учтем, что $r_k(x)=d(x)$, и положим $u_1(x)=1$, $v_1(x)=-q_k(x)$, то предпоследнее из равенств (2) даст:

$$d(x)=r_{k-2}(x)u_1(x)+r_{k-1}(x)v_1(x).$$

Подставляя сюда выражение $r_{k-1}(x)$ через $r_{k-3}(x)$ и $r_{k-2}(x)$ из предшествующего равенства (2), мы получим:

$$d(x)=r_{k-3}(x)u_2(x)+r_{k-2}(x)v_2(x),$$

где, очевидно, $u_2(x)=v_1(x)$, $v_2(x)=u_1(x)-v_1(x)q_{k-1}(x)$. Продолжая подниматься вверх по равенствам (2), мы придем, наконец, к доказываемому равенству (3).

Для доказательства второго утверждения теоремы предположим, что многочлены $u(x)$ и $v(x)$, удовлетворяющие равенству (3), уже

найдены, но, например, степень $u(x)$ больше или равна степени $g(x)$. Делим $u(x)$ на $g(x)$:

$$u(x) = g(x)\,q(x) + r(x),$$

где степень $r(x)$ меньше степени $g(x)$, и подставляем это выражение в (3). Мы получим равенство

$$f(x)\,r(x) + g(x)\,[v(x) + f(x)\,q(x)] = d(x).$$

Степень множителя, стоящего при $f(x)$, уже меньше степени $g(x)$. Степень многочлена, стоящего в квадратных скобках, будет в свою очередь меньше степени $f(x)$, так как в противном случае степень второго слагаемого левой части была бы не меньше степени произведения $g(x)f(x)$, а так как степень первого слагаемого меньше степени этого произведения, то вся левая часть имела бы степень, бóльшую или равную степени $g(x)f(x)$, тогда как многочлен $d(x)$ заведомо имеет, при наших предположениях, меньшую степень.

Теорема доказана. Одновременно мы получаем, что если многочлены $f(x)$ и $g(x)$ имеют рациональные или действительные коэффициенты, то и многочлены $u(x)$ и $v(x)$, удовлетворяющие равенству (3), можно подобрать так, что их коэффициенты будут рациональными или, соответственно, действительными.

Пример. Найдем многочлены $u(x)$ и $v(x)$, удовлетворяющие равенству (3) при

$$f(x) = x^3 - x^2 + 3x - 10, \quad g(x) = x^3 + 6x^2 - 9x - 14.$$

Применим к этим многочленам алгоритм Евклида, причем теперь при выполнении делений уже нельзя допускать искажения частных, так как эти частные используются при разыскании многочленов $u(x)$ и $v(x)$. Мы получим такую систему равенств:

$$f(x) = g(x) + (-7x^2 + 12x + 4);$$

$$g(x) = (-7x^2 + 12x + 4)\left(-\frac{1}{7}x - \frac{54}{49}\right) + \frac{235}{49}(x-2);$$

$$-7x^2 + 12x + 4 = (x-2)(-7x-2).$$

Отсюда следует, что $(f(x), g(x)) = x - 2$ и что

$$u(x) = \frac{7}{235}x + \frac{54}{235}, \quad v(x) = -\frac{7}{235}x - \frac{5}{235}.$$

Применяя доказанную сейчас теорему к взаимно простым многочленам, мы получаем такой результат:

Многочлены $f(x)$ и $g(x)$ тогда и только тогда взаимно просты, если можно найти многочлены $u(x)$ и $v(x)$, удовлетворяющие равенству

$$f(x)\,u(x) + g(x)\,v(x) = 1. \tag{4}$$

Опираясь на этот результат, можно доказать несколько простых, но важных теорем о взаимно простых многочленах:

а) *Если многочлен $f(x)$ взаимно прост с каждым из многочленов $\varphi(x)$ и $\psi(x)$, то он взаимно прост и с их произведением.*

В самом деле, существуют, по (4), такие многочлены $u(x)$ и $v(x)$, что

$$f(x)\,u(x)+\varphi(x)\,v(x)=1.$$

Умножая это равенство на $\psi(x)$, получаем:

$$f(x)\,[u(x)\,\psi(x)]+[\varphi(x)\,\psi(x)]\,v(x)=\psi(x),$$

откуда следует, что всякий общий делитель $f(x)$ и $\varphi(x)\,\psi(x)$ был бы делителем и для $\psi(x)$; однако по условию $(f(x),\ \psi(x))=1$.

б) *Если произведение многочленов $f(x)$ и $g(x)$ делится на $\varphi(x)$, но $f(x)$ и $\varphi(x)$ взаимно просты, то $g(x)$ делится на $\varphi(x)$.*

В самом деле, умножая равенство

$$f(x)\,u(x)+\varphi(x)\,v(x)=1$$

на $g(x)$, мы получим:

$$[f(x)\,g(x)]\,u(x)+\varphi(x)\,[v(x)\,g(x)]=g(x).$$

Оба слагаемых левой части этого равенства делятся на $\varphi(x)$; на него делится, следовательно, и $g(x)$.

в) *Если многочлен $f(x)$ делится на каждый из многочленов $\varphi(x)$ и $\psi(x)$, которые между собой взаимно просты, то $f(x)$ делится и на их произведение.*

Действительно, $f(x)=\varphi(x)\,\overline{\varphi}(x)$, так что произведение, стоящее справа, делится на $\psi(x)$. Поэтому, по б), $\overline{\varphi}(x)$ делится на $\psi(x)$, $\overline{\varphi}(x)=\psi(x)\,\overline{\psi}(x)$, откуда $f(x)=[\varphi(x)\,\psi(x)]\,\overline{\psi}(x)$.

Определение наибольшего общего делителя может быть распространено на случай любой конечной системы многочленов: *наибольшим общим делителем* многочленов $f_1(x),\ f_2(x),\ \ldots,\ f_s(x)$ называется такой общий делитель этих многочленов, который делится на любой другой общий делитель этих многочленов. Существование наибольшего общего делителя для любой конечной системы многочленов вытекает из следующей теоремы, дающей также способ его вычисления.

Наибольший общий делитель многочленов $f_1(x),\ f_2(x),\ \ldots,\ f_s(x)$ равен наибольшему общему делителю многочлена $f_s(x)$ и наибольшего общего делителя многочленов $f_1(x),\ f_2(x),\ \ldots,\ f_{s-1}(x)$.

В самом деле, при $s=2$ теорема очевидна. Мы примем поэтому, что для случая $s-1$ она справедлива, т. е., в частности, уже доказано существование наибольшего общего делителя $d(x)$ многочленов $f_1(x),\ f_2(x),\ \ldots,\ f_{s-1}(x)$. Обозначим через $\overline{d}(x)$ наибольший общий делитель многочленов $d(x)$ и $f_s(x)$. Он будет, очевидно, общим делителем для всех заданных многочленов. С другой стороны, всякий другой общий делитель этих многочленов будет делителем также и для $d(x)$, а поэтому и для $\overline{d}(x)$.

В частности, система многочленов $f_1(x)$, $f_2(x)$, ..., $f_s(x)$ называется *взаимно простой*, если общими делителями этих многочленов являются лишь многочлены нулевой степени, т. е. если их наибольший общий делитель равен 1. Если $s>2$, то попарно эти многочлены могут и не быть взаимно простыми. Так, система многочленов

$$f(x)=x^3-7x^2+7x+15, \quad g(x)=x^2-x-20,$$
$$h(x)=x^3+x^2-12x$$

взаимно проста, хотя

$$(f(x),\ g(x))=x-5,\ (f(x),\ h(x))=x-3,\ (g(x),\ h(x))=x+4.$$

Читатель без труда получит обобщение доказанных выше теорем а) — в) о взаимно простых многочленах на случай любого конечного числа многочленов.

§ 22. Корни многочленов

Мы уже встречались в § 20 со значениями многочлена, когда говорили о теоретико-функциональной точке зрения на понятие многочлена. Напомним определение.

Если

$$f(x)=a_0x^n+a_1x^{n-1}+\ \dots\ +a_n \qquad (1)$$

есть некоторый многочлен, а c — некоторое число, то число

$$f(c)=a_0c^n+a_1c^{n-1}+\ \dots\ +a_n,$$

полученное заменой в выражении (1) для $f(x)$ неизвестного x числом c и последующим выполнением всех указанных операций, называется *значением многочлена* $f(x)$ *при* $x=c$. Понятно, что если $f(x)=g(x)$ в смысле алгебраического равенства многочленов, определенного в § 20, то $f(c)=g(c)$ при любом c.

Легко видеть также, что если

$$\varphi(x)=f(x)+g(x),\quad \psi(x)=f(x)g(x),$$

то

$$\varphi(c)=f(c)+g(c),\quad \psi(c)=f(c)g(c).$$

Иными словами, сложение и умножение многочленов, определенные в § 20, превращаются при теоретико-функциональной точке зрения на многочлены в сложение и умножение функций, понимаемые в смысле сложения или умножения соответственных значений этих функций.

Если $f(c)=0$, т. е. многочлен $f(x)$ обращается в нуль при подстановке в него числа c вместо неизвестного, то c называется *корнем* многочлена $f(x)$ (или уравнения $f(x)=0$). Сейчас будет показано, что это понятие целиком относится к той теории делимости многочленов, которая была предметом изучения в предшествующем параграфе.

Если мы будем делить многочлен $f(x)$ на произвольный многочлен первой степени (или, как будем говорить дальше, на *линейный многочлен*), то остаток будет либо некоторым многочленом нулевой степени, либо нулем, т. е. во всяком случае некоторым числом r. Следующая теорема позволяет найти этот остаток, не выполняя самого деления, в случае, когда производится деление на многочлен вида $x-c$.

Остаток от деления многочлена $f(x)$ на линейный многочлен $x-c$ равен значению $f(c)$ многочлена $f(x)$ при $x=c$.

Действительно, пусть

$$f(x)=(x-c)\,q(x)+r.$$

Беря значения обеих частей этого равенства при $x=c$, мы получаем:

$$f(c)=(c-c)\,q(c)+r=r,$$

что доказывает теорему.

Отсюда вытекает такое исключительно важное следствие:

Число c тогда и только тогда будет корнем многочлена $f(x)$, если $f(x)$ делится на $x-c$.

С другой стороны, если $f(x)$ делится на некоторый многочлен первой степени $ax+b$, то делится, очевидно, и на многочлен $x-\left(-\frac{b}{a}\right)$, т. е. на многочлен вида $x-c$. Таким образом, *разыскание корней многочлена $f(x)$ равносильно разысканию его линейных делителей.*

Ввиду сказанного выше представляет интерес следующий метод деления многочлена $f(x)$ на линейный двучлен $x-c$, более простой, чем общий алгоритм деления многочленов. Этот метод называется методом Горнера. Пусть

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n \tag{2}$$

и пусть

$$f(x)=(x-c)\,q(x)+r, \tag{3}$$

где

$$q(x)=b_0x^{n-1}+b_1x^{n-2}+b_2x^{n-3}+\ldots+b_{n-1}.$$

Сравнивая коэффициенты при одинаковых степенях x в (3), получаем:

$$\begin{aligned}
a_0&=b_0,\\
a_1&=b_1-cb_0,\\
a_2&=b_2-cb_1,\\
&\ldots\ldots\\
a_{n-1}&=b_{n-1}-cb_{n-2},\\
a_n&=r-cb_{n-1}.
\end{aligned}$$

Отсюда следует, что $b_0=a_0$, $b_k=cb_{k-1}+a_k$, $k=1, 2, \ldots, n-1$, т. е. коэффициент b_k получается умножением преды-

дущего коэффициента b_{k-1} на c и прибавлением соответствующего коэффициента a_k; наконец, $r=cb_{n-1}+a_n$, т. е. и остаток r, равный, как мы знаем, $f(c)$, получается по этому же закону. Таким образом, коэффициенты частного и остаток можно последовательно получать при помощи однотипных вычислений, которые располагаются в схему, как показывают следующие примеры:

1. Разделить $f(x)=2x^5-x^4-3x^3+x-3$ на $x-3$.

Составим таблицу, в которой над чертой расположены коэффициенты многочлена $f(x)$, под чертой — соответствующие коэффициенты частного и остаток, последовательно вычисляемые, а слева сбоку — значение c в данном примере:

$$\begin{array}{c|cccccc} & 2 & -1 & -3 & 0 & 1 & -3 \\ \hline 3 & 2, & 3\cdot 2-1=5, & 3\cdot 5-3=12, & 3\cdot 12+0=36, & 3\cdot 36+1=109, & 3\cdot 109-3=324 \end{array}.$$

Таким образом, искомое частное будет

$$q(x)=2x^4+5x^3+12x^2+36x+109,$$

а остаток $r=f(3)=324$.

2. Разделить $f(x)=x^4-8x^3+x^2+4x-9$ на $x+1$.

$$\begin{array}{c|ccccc} & 1 & -8 & 1 & 4 & -9 \\ \hline -1 & 1 & -9 & 10 & -6 & -3. \end{array}$$

Поэтому частное будет

$$q(x)=x^3-9x^2+10x-6,$$

а остаток $r=f(-1)=-3$.

Эти примеры показывают, что *метод Горнера может быть использован также для быстрого вычисления значения многочлена при данном значении неизвестного.*

Кратные корни. Если c — корень многочлена $f(x)$, т. е. $f(c)=0$, то $f(x)$ делится, как мы знаем, на $x-c$. Может оказаться, что многочлен $f(x)$ делится не только на первую степень линейного двучлена $x-c$, но и на более высокие его степени. Во всяком случае найдется такое натуральное число k, что $f(x)$ нацело делится на $(x-c)^k$, но не делится на $(x-c)^{k+1}$. Поэтому

$$f(x)=(x-c)^k\varphi(x),$$

где многочлен $\varphi(x)$ на $x-c$ уже не делится, т. е. число c своим корнем не имеет. Число k называется *кратностью* корня c в многочлене $f(x)$, а сам корень c — *k-кратным корнем* этого многочлена. Если $k=1$, то говорят, что корень c — *простой.*

Понятие кратного корня тесно связано с понятием производной от многочлена. Мы изучаем, однако, многочлены с любыми комплексными коэффициентами и поэтому не можем просто воспользоваться понятием производной, введенным в курсе математического анализа. То, что будет сказано ниже, следует рассматривать как

независимое от курса анализа определение производной многочлена.

Пусть дан многочлен n-й степени

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

с любыми комплексными коэффициентами. Его *производной* (или *первой производной*) называется многочлен $(n-1)$-й степени

$$f'(x)=na_0x^{n-1}+(n-1)a_1x^{n-2}+\ldots+2a_{n-2}x+a_{n-1}.$$

Производная от многочлена нулевой степени и от нуля считается равной нулю. Производная от первой производной называется *второй производной* от многочлена $f(x)$ и обозначается через $f''(x)$ и т. д. Очевидно, что

$$f^{(n)}(x)=n!a_0$$

и поэтому $f^{(n+1)}(x)=0$, т. е. $(n+1)$-*я производная от многочлена* n-*й степени равна нулю.*

Мы не можем пользоваться в нашем случае многочленов с комплексными коэффициентами свойствами производной, доказанными в курсе анализа для многочленов с действительными коэффициентами, и должны, используя лишь данное выше определение производной, снова эти свойства доказать. Нас интересуют следующие свойства, являющиеся, как говорят, формулами дифференцирования для суммы и произведения:

$$(f(x)+g(x))'=f'(x)+g'(x), \qquad (4)$$

$$(f(x)\cdot g(x))'=f(x)g'(x)+f'(x)g(x). \qquad (5)$$

Эти формулы легко проверить, впрочем, непосредственным подсчетом, беря в качестве $f(x)$ и $g(x)$ два произвольных многочлена и применяя данное выше определение производной; эту проверку мы предоставим читателю.

Формула (5) без труда распространяется на случай произведения любого конечного числа множителей, а поэтому обычным способом может быть выведена формула и для производной от степени:

$$(f^k(x))'=kf^{k-1}(x)f'(x). \qquad (6)$$

Нашей целью является доказательство следующей теоремы:

Если число c является k-кратным корнем многочлена $f(x)$, то при $k>1$ оно будет $(k-1)$-кратным корнем первой производной этого многочлена; если же $k=1$, то c не будет служить корнем для $f'(x)$.

В самом деле, пусть

$$f(x)=(x-c)^k\varphi(x), \quad k\geqslant 1, \qquad (7)$$

где $\varphi(x)$ уже не делится на $x-c$. Дифференцируя равенство (7), получаем:

$$f'(x)=(x-c)^k\varphi'(x)+k(x-c)^{k-1}\varphi(x)=$$
$$=(x-c)^{k-1}[(x-c)\varphi'(x)+k\varphi(x)].$$

Первое слагаемое суммы, стоящей в квадратных скобках, делится на $x-c$, а второе на $x-c$ не делится; поэтому вся эта сумма на $x-c$ не может делиться. Учитывая, что частное от деления $f(x)$ на $(x-c)^{k-1}$ определено однозначно, мы получаем, что $(x-c)^{k-1}$ является наибольшей степенью двучлена $x-c$, на которую делится многочлен $f'(x)$, что и требовалось доказать.

Применяя эту теорему несколько раз, мы получаем, что *k-кратный корень многочлена $f(x)$ будет $(k-s)$-кратным в s-й производной этого многочлена $(k \geqslant s)$ и впервые не будет служить корнем для k-й производной от $f(x)$.*

§ 23. Основная теорема

Занимаясь в предшествующем параграфе корнями многочленов, мы не ставили вопроса о том, всякий ли многочлен обладает корнями. Известно, что существуют многочлены с действительными коэффициентами, не имеющие действительных корней; x^2+1 — один из таких многочленов. Можно было бы ожидать, что существуют многочлены, не имеющие корней даже среди комплексных чисел, особенно если рассматриваются многочлены с любыми комплексными коэффициентами. Если бы это было так, то система комплексных чисел нуждалась бы в дальнейшем расширении. На самом деле, однако, справедлива следующая основная теорема алгебры комплексных чисел:

Всякий многочлен с любыми числовыми коэффициентами, степень которого не меньше единицы, имеет хотя бы один корень, в общем случае комплексный.

Эта теорема является одним из крупнейших достижений всей математики и находит применения в самых различных областях науки. На ней основана, в частности, вся дальнейшая теория многочленов с числовыми коэффициентами, и потому эту теорему называли раньше (а иногда называют и теперь) «основной теоремой высшей алгебры». В действительности, однако, основная теорема не является чисто алгебраической. Все ее доказательства,— а их, после Гаусса, впервые доказавшего эту теорему в самом конце XVIII века, было найдено весьма много,— принуждены в большей или меньшей мере использовать так называемые топологические свойства действительных и комплексных чисел, т. е. свойства, связанные с непрерывностью.

В доказательстве, которое будет сейчас проведено, многочлен $f(x)$ с комплексными коэффициентами будет рассматриваться как

комплексная функция комплексного переменного x. Таким образом, x может принимать любые комплексные значения, т. е., как говорят, учитывая изложенный в § 17 способ построения комплексных чисел, переменное x изменяется на *комплексной плоскости*. Значения функции $f(x)$ также будут комплексными числами. Можно считать, что эти значения отмечаются на втором экземпляре комплексной плоскости, подобно тому как в случае действительных функций действительного переменного значения независимого переменного отмечаются на одной числовой прямой (оси абсцисс), а значения функции — на другой (оси ординат).

Определение непрерывной функции, известное читателю из курса математического анализа, переносится и на функции комплексного переменного, причем в формулировке определения абсолютные величины заменяются модулями.

Именно, комплексная функция $f(x)$ комплексного переменного x называется *непрерывной в точке* x_0, если для всякого положительного действительного числа ε можно подобрать такое положительное действительное число δ, что, каково бы ни было (вообще говоря, комплексное) приращение h, модуль которого удовлетворяет неравенству $|h|<\delta$, будет справедливым также неравенство

$$|f(x_0+h)-f(x_0)|<\varepsilon.$$

Функция $f(x)$ называется *непрерывной*, если она непрерывна во всех точках x_0, в которых она определена, т. е. если $f(x)$ является многочленом, — на всей комплексной плоскости.

Многочлен $f(x)$ является непрерывной функцией комплексного переменного x.

Доказательство этой теоремы можно было бы провести так же, как это делается в курсе математического анализа, а именно, показав, что сумма и произведение непрерывных функций сами непрерывны, и заметив, что функция, постоянно равная одному и тому же комплексному числу, будет непрерывной. Мы пойдем, однако, иным путем.

Докажем сначала частный случай теоремы, а именно случай, когда свободный член многочлена $f(x)$ равен нулю, причем докажем лишь непрерывность $f(x)$ в точке $x_0=0$. Иными словами, мы докажем следующую лемму (вместо h мы пишем x):

Лемма 1. *Если свободный член многочлена $f(x)$ равен нулю:*

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x,$$

т. е. $f(0)=0$, то для всякого $\varepsilon>0$ можно подобрать такое $\delta>0$, что при всех x, для которых $|x|<\delta$, будет $|f(x)|<\varepsilon$.

Действительно, пусть

$$A=\max(|a_0|, |a_1|, \ldots, |a_{n-1}|).$$

Число ε нам уже дано. Покажем, что если за число δ взять

$$\delta=\frac{\varepsilon}{A+\varepsilon}, \tag{1}$$

то оно будет удовлетворять требуемым условиям.

В самом деле,

$$|f(x)|\leqslant|a_0||x|^n+|a_1||x|^{n-1}+\ldots+|a_{n-1}||x|\leqslant$$
$$\leqslant A(|x|^n+|x|^{n-1}+\ldots+|x|),$$

т. е.

$$|f(x)|\leqslant A\frac{|x|-|x|^{n+1}}{1-|x|}.$$

Так как $|x|<\delta$ и, по (1), $\delta<1$, то

$$\frac{|x|-|x|^{n+1}}{1-|x|}<\frac{|x|}{1-|x|},$$

и поэтому

$$|f(x)|<\frac{A|x|}{1-|x|}<\frac{A\delta}{1-\delta}=\frac{A\frac{\varepsilon}{A+\varepsilon}}{1-\frac{\varepsilon}{A+\varepsilon}}=\varepsilon,$$

что и требовалось доказать.

Выведем теперь следующую формулу. Пусть дан многочлен

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

с любыми комплексными коэффициентами. Подставим в него вместо x сумму $x+h$, где h—второе неизвестное. Разлагая в правой части каждую из степеней $(x+h)^k$, $k\leqslant n$, по формуле бинома и собирая вместе члены с одинаковыми степенями h, мы получим, как читатель без труда проверит, равенство

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^n}{n!}f^{(n)}(x),$$

т. е. докажем *формулу Тэйлора*, дающую разложение $f(x+h)$ по степеням «приращения» h.

Непрерывность произвольного многочлена $f(x)$ в любой точке x_0 доказывается теперь следующим образом. По формуле Тэйлора

$$f(x_0+h)-f(x_0)=c_1h+c_2h^2+\ldots+c_nh^n=\varphi(h),$$

где

$$c_1=f'(x_0),\quad c_2=\frac{1}{2!}f''(x_0),\quad\ldots,\quad c_n=\frac{1}{n!}f^{(n)}(x_0).$$

Многочлен $\varphi(h)$ от неизвестного h есть многочлен без свободного члена, поэтому, по лемме 1, для всякого $\varepsilon > 0$ можно подобрать такое $\delta > 0$, что при $|h| < \delta$ будет $|\varphi(h)| < \varepsilon$, т. е.

$$|f(x_0+h)-f(x_0)| < \varepsilon,$$

что и требовалось доказать.

Из неравенства

$$||f(x_0+h)|-|f(x_0)|| \leqslant |f(x_0+h)-f(x_0)|,$$

основанного на формуле (13) § 18, и из доказанной сейчас непрерывности многочлена вытекает *непрерывность модуля* $|f(x)|$ многочлена $f(x)$; этот модуль является, очевидно, действительной неотрицательной функцией комплексного переменного x.

Сейчас будут доказаны леммы, используемые при доказательстве основной теоремы.

Лемма о модуле старшего члена. *Если дан многочлен n-й степени, $n \geqslant 1$,*

$$f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n$$

с произвольными комплексными коэффициентами и если k — любое положительное действительное число, то для достаточно больших по модулю значений неизвестного x имеет место неравенство

$$|a_0x^n| > k|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n|, \qquad (2)$$

т. е. модуль старшего члена будет больше модуля суммы всех остальных членов, притом во сколько угодно раз.

В самом деле, пусть A — наибольший из модулей коэффициентов $a_1, a_2, \ldots, a_n$:

$$A = \max(|a_1|, |a_2|, \ldots, |a_n|).$$

Тогда (см. в § 18 свойства модулей суммы и произведения комплексных чисел)

$$|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n| \leqslant |a_1||x|^{n-1} + |a_2||x|^{n-2} + \ldots$$
$$\ldots + |a_n| \leqslant A(|x|^{n-1} + |x|^{n-2} + \ldots + 1) = A\frac{|x|^n - 1}{|x|-1}.$$

Полагая $|x| > 1$, мы получим:

$$\frac{|x|^n-1}{|x|-1} < \frac{|x|^n}{|x|-1},$$

откуда

$$|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n| < A\frac{|x|^n}{|x|-1}.$$

Таким образом, неравенство (2) будет выполняться, если x удовлетворяет, помимо условия $|x|>1$, также неравенству

$$kA\frac{|x|^n}{|x|-1} \leqslant |a_0x^n| = |a_0||x|^n,$$

т. е. если

$$|x| \geqslant \frac{kA}{|a_0|}+1. \tag{3}$$

Так как правая часть неравенства (3) больше 1, то можно утверждать, что для значений x, удовлетворяющих этому неравенству, имеет место неравенство (2), что доказывает лемму.

Лемма о возрастании модуля многочлена. *Для всякого многочлена $f(x)$ с комплексными коэффициентами, степень которого не меньше единицы, и всякого положительного действительного числа M, сколь угодно большого, можно подобрать такое положительное действительное число N, что при $|x|>N$ будет $|f(x)|>M$.*

Пусть

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n.$$

По формуле (11) § 18

$$|f(x)| = |a_0x^n + (a_1x^{n-1} + \ldots + a_n)| \geqslant$$
$$\geqslant |a_0x^n| - |a_1x^{n-1} + \ldots + a_n|. \tag{4}$$

Применим лемму о модуле старшего члена, положив $k=2$: существует такое число N_1, что при $|x|>N_1$ будет

$$|a_0x^n| > 2|a_1x^{n-1} + \ldots + a_n|.$$

Отсюда

$$|a_1x^{n-1} + \ldots + a_n| < \frac{1}{2}|a_0x^n|,$$

т. е., по (4),

$$|f(x)| > |a_0x^n| - \frac{1}{2}|a_0x^n| = \frac{1}{2}|a_0x^n|.$$

Правая часть этого неравенства будет больше M при

$$|x| > N_2 = \sqrt[n]{\frac{2M}{|a_0|}}.$$

Таким образом, при $|x|>N=\max(N_1, N_2)$ будет $|f(x)|>M$.

Смысл этой леммы может быть выяснен при помощи следующей геометрической иллюстрации, которая в настоящем параграфе будет неоднократно использоваться. Предположим, что в каждой точке x_0 комплексной плоскости к этой плоскости восставлен перпендикуляр, длина которого (при заданной единице масштаба) равна модулю значения многочлена $f(x)$ в этой точке, т. е. равна $|f(x_0)|$. Концы перпендикуляров будут составлять ввиду доказанной выше

непрерывности модуля многочлена некоторую непрерывную кривую поверхность, расположенную над комплексной плоскостью. Лемма о возрастании модуля многочлена показывает, что эта поверхность при возрастании $|x_0|$ все больше и больше удаляется от комплексной плоскости, хотя, понятно, это удаление вовсе не является монотонным. Рис. 8 схематически изображает линию пересечения этой поверхности с плоскостью, перпендикулярной к комплексной плоскости и проходящей через точку O.

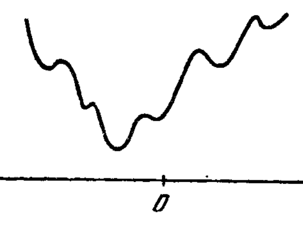


Рис. 8.

Основную роль в доказательстве играет следующая лемма:

Лемма Даламбера. *Если при $x = x_0$ многочлен $f(x)$ степени n, $n \geqslant 1$, не обращается в нуль, $f(x_0) \neq 0$ и поэтому $|f(x_0)| > 0$, то можно найти такое приращение h, в общем случае комплексное, что*

$$|f(x_0 + h)| < |f(x_0)|.$$

По формуле Тэйлора, если приращение h пока произвольно, будет

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \ldots + \frac{h^n}{n!} f^{(n)}(x_0).$$

По условию, x_0 не является корнем для $f(x)$. Случайно, однако, это число может оказаться корнем для $f'(x)$, а также, быть может, для некоторых из дальнейших производных. Пусть k-я производная ($k \geqslant 1$) будет первой, не имеющей x_0 своим корнем, т. е.

$$f'(x_0) = f''(x_0) = \ldots = f^{(k-1)}(x_0) = 0, \quad f^{(k)}(x_0) \neq 0.$$

Такое k существует, так как, если a_0 есть старший коэффициент многочлена $f(x)$, то

$$f^{(n)}(x_0) = n! a_0 \neq 0.$$

Таким образом,

$$f(x_0 + h) = f(x_0) + \frac{h^k}{k!} f^{(k)}(x_0) + \\ + \frac{h^{k+1}}{(k+1)!} f^{(k+1)}(x_0) + \ldots + \frac{h^n}{n!} f^{(n)}(x_0).$$

Некоторые из чисел $f^{(k+1)}(x_0), \ldots, f^{(n-1)}(x_0)$ также могут равняться нулю, но это для нас несущественно.

Деля обе части этого равенства на $f(x_0)$, отличное, по условию, от нуля, и вводя обозначение

$$c_j = \frac{f^{(j)}(x_0)}{j!\, f(x_0)}, \quad j = k,\ k+1,\ \ldots,\ n,$$

мы получим:

$$\frac{f(x_0+h)}{f(x_0)} = 1 + c_k h^k + c_{k+1} h^{k+1} + \ldots + c_n h^n$$

или, ввиду $c_k \neq 0$,

$$\frac{f(x_0+h)}{f(x_0)} = (1 + c_k h^k) + c_k h^k \left(\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k}\right).$$

Переходя к модулям, получим:

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| \leqslant |1 + c_k h^k| + |c_k h^k| \left|\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k}\right|. \quad (5)$$

До этого момента мы не делали никаких предположений о приращении h. Теперь мы будем выбирать h, причем будем отдельно выбирать его модуль и его аргумент. Модуль h будет выбираться следующим образом. Так как

$$\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k}$$

является многочленом от h без свободного члена, то, по лемме 1 $\left(\text{полагая } \varepsilon = \frac{1}{2}\right)$, можно найти такое δ_1, что при $|h| < \delta_1$ будет

$$\left|\frac{c_{k+1}}{c_k} h + \ldots + \frac{c_n}{c_k} h^{n-k}\right| < \frac{1}{2}. \quad (6)$$

С другой стороны, при

$$|h| < \delta_2 = \sqrt[k]{|c_k|^{-1}}$$

будет

$$|c_k h^k| < 1. \quad (7)$$

Положим, что модуль h выбран в соответствии с неравенством

$$|h| < \min(\delta_1, \delta_2). \quad (8)$$

Тогда, ввиду (6), неравенство (5) превращается в строгое неравенство

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| < |1 + c_k h^k| + \frac{1}{2} |c_k h^k|; \quad (9)$$

условием (7) мы воспользуемся лишь позже.

Для выбора аргумента h потребуем, чтобы число $c_k h^k$ было отрицательным действительным числом. Иными словами,

$$\arg(c_k h^k) = \arg c_k + k \arg h = \pi,$$

откуда

$$\arg h = \frac{\pi - \arg c_k}{k}. \quad (10)$$

При этом выборе h число $c_k h^k$ будет отличаться знаком от своей абсолютной величины,

$$c_k h^k = -|c_k h^k|,$$

а поэтому, используя неравенство (7),

$$|1 + c_k h^k| = |1 - |c_k h^k|| = 1 - |c_k h^k|.$$

Таким образом, при выборе h на основании условий (8) и (10) неравенство (9) принимает вид

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| < 1 - |c_k h^k| + \frac{1}{2}|c_k h^k| = 1 - \frac{1}{2}|c_k h^k|,$$

т. е. тем более

$$\left|\frac{f(x_0+h)}{f(x_0)}\right| = \frac{|f(x_0+h)|}{|f(x_0)|} < 1,$$

откуда следует

$$|f(x_0+h)| < |f(x_0)|,$$

что доказывает лемму Даламбера.

При помощи той геометрической иллюстрации, которая была дана выше, можно следующим образом пояснить лемму Даламбера. Дано, что $|f(x_0)| > 0$. Это значит, что длина перпендикуляра, восставленного к комплексной плоскости в точке x_0, отлична от нуля. Тогда, по лемме Даламбера, можно найти такую точку $x_1 = x_0 + h$, что $|f(x_1)| < |f(x_0)|$, т. е. перпендикуляр в точке x_1 будет более коротким, чем в точке x_0, и, следовательно, поверхность, образованная концами перпендикуляров, будет в этой новой точке несколько ближе к комплексной плоскости. Как показывает доказательство леммы, модуль h можно считать сколь угодно ма́лым, т. е. точку x_1 можно выбрать как угодно близко к точке x_0; мы не будем, однако, пользоваться в дальнейшем этим замечанием.

Корнями многочлена $f(x)$ будут служить, очевидно, те комплексные числа (т. е. те точки комплексной плоскости), в которых поверхность, образованная концами перпендикуляров, коснется этой плоскости. Опираясь лишь на лемму Даламбера, нельзя доказать существование таких точек. В самом деле, пользуясь этой леммой, можно найти такую бесконечную последовательность точек $x_0, x_1, x_2, \ldots$, что

$$|f(x_0)| > |f(x_1)| > |f(x_2)| > \ldots, \tag{11}$$

Отсюда не следует, однако, существование такой точки $\bar{x}$, что $f(\bar{x}) = 0$, тем более, что убывающая последовательность положительных действительных чисел (11) вовсе не обязана стремиться к нулю.

Дальнейшие рассмотрения будут основаны на одной теореме из теории функций комплексного переменного, обобщающей теорему Вейерштрасса, известную читателю из курса математического анализа. Она относится к действительным функциям комплексного перемен-

ного, т. е. к функциям комплексного переменного, принимающим лишь действительные значения; примером таких функций служит модуль многочлена. В формулировке этой теоремы мы будем говорить для простоты *о замкнутом круге* E, понимая под этим круг на комплексной плоскости, к которому присоединены все точки его границы.

Если действительная функция $g(x)$ комплексного переменного x непрерывна во всех точках замкнутого круга E, то существует в круге E такая точка x_0, что для всех x из E имеет место неравенство $g(x) \geqslant g(x_0)$. Точка x_0 является, следовательно, точкой минимума для $g(x)$ в круге E.

Доказательство этой теоремы можно найти во всех курсах теории функций комплексного переменного, и мы его не приводим

Ограничиваясь случаем, когда функция $g(x)$ неотрицательна во всех точках круга E,— только этот случай представляет для нас интерес,— поясним геометрически эту теорему при помощи той иллюстрации, которая уже использована выше. В каждой точке x_0 круга E проводим перпендикуляр длины $g(x_0)$. Концы этих перпендикуляров составляют кусок непрерывной кривой поверхности, причем благодаря замкнутости круга E существование точек минимума для этого куска поверхности делается геометрически достаточно ясным. Эта иллюстрация не заменяет, конечно, доказательства теоремы.

Теперь мы можем перейти к непосредственному доказательству основной теоремы. Пусть дан многочлен $f(x)$ степени n, $n \geqslant 1$. Если его свободный член есть a_n, то, очевидно, $f(0) = a_n$. Применим к нашему многочлену лемму о возрастании модуля многочлена, полагая $M = |f(0)| = |a_n|$. Существует, следовательно, такое N, что при $|x| > N$ будет $|f(x)| > |f(0)|$. Очевидно, далее, что указанное выше обобщение теоремы Вейерштрасса применимо к функции $|f(x)|$ при любом выборе замкнутого круга E. В качестве E мы возьмем замкнутый круг, ограниченный окружностью радиуса N с центром в точке 0. Пусть точка x_0 будет точкой минимума для $|f(x)|$ в круге E, откуда, в частности, следует $|f(x_0)| \leqslant |f(0)|$.

Легко видеть, что x_0 *на самом деле будет служить точкой минимума для $|f(x)|$ на всей комплексной плоскости*: если точка x' лежит вне E, то $|x'| > N$, и поэтому

$$|f(x')| > |f(0)| \geqslant |f(x_0)|.$$

Отсюда следует, наконец, что $f(x_0) = 0$, т. е. что x_0 *служит корнем для* $f(x)$; если бы было $f(x_0) \neq 0$, то, по лемме Даламбера, существовала бы такая точка x_1, что $|f(x_1)| < |f(x_0)|$; это противоречит, однако, только что установленному свойству точки x_0.

Заметим, что еще одно доказательство основной теоремы будет приведено в § 55.

§ 24. Следствия из основной теоремы

Пусть дан многочлен n-й степени, $n \geqslant 1$,

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n \tag{1}$$

с любыми комплексными коэффициентами. Мы снова рассматриваем его как формально-алгебраическое выражение, вполне определяемое набором своих коэффициентов. Основная теорема о существовании корня, доказанная в предшествующем параграфе, позволяет утверждать существование для $f(x)$ корня α_1, комплексного или действительного. Поэтому многочлен $f(x)$ обладает разложением

$$f(x) = (x - \alpha_1)\, \varphi(x).$$

Коэффициенты многочлена $\varphi(x)$ снова являются действительными или комплексными числами, и поэтому $\varphi(x)$ обладает корнем α_2, откуда

$$f(x) = (x - \alpha_1)(x - \alpha_2)\, \psi(x).$$

Продолжая так далее, мы придем после конечного числа шагов к *разложению многочлена n-й степени $f(x)$ в произведение n линейных множителей*,

$$f(x) = a_0 (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n). \tag{2}$$

Коэффициент a_0 появился по следующей причине: если бы справа в выражении (2) стоял некоторый коэффициент b, то после раскрытия скобок старший член многочлена $f(x)$ имел бы вид bx^n, хотя на самом деле, ввиду (1), им является член $a_0 x^n$. Поэтому $b = a_0$.

Разложение (2) *является для многочлена $f(x)$ единственным с точностью до порядка сомножителей разложением такого типа.*

Пусть, в самом деле, имеется еще разложение

$$f(x) = a_0 (x - \beta_1)(x - \beta_2) \ldots (x - \beta_n). \tag{3}$$

Из (2) и (3) следует равенство

$$(x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n) = (x - \beta_1)(x - \beta_2) \ldots (x - \beta_n). \tag{4}$$

Если бы корень α_i был отличен от всех β_j, $j = 1, 2, \ldots, n$, то, подставляя α_i вместо неизвестного в (4), мы получили бы слева нуль, а справа число, отличное от нуля. Таким образом, *всякий корень α_i равен некоторому корню β_j и обратно.*

Отсюда еще не вытекает совпадение разложений (2) и (3). Действительно, среди корней α_i, $i = 1, 2, \ldots, n$, могут быть равные между собой. Пусть, например, s этих корней равны α_1 и пусть, с другой стороны, среди корней β_j, $j = 1, 2, \ldots, n$, содержится t равных корню α_1. Нужно показать, что $s = t$.

Так как степень произведения многочленов равна сумме степеней сомножителей, то произведение двух многочленов, отличных от нуля, не может равняться нулю. Отсюда вытекает, что если два произведения многочленов равны друг другу, *то обе части равенства можно сократить на общий множитель*: если

$$f(x)\,\varphi(x) = g(x)\,\varphi(x)$$

и $\varphi(x) \neq 0$, то из

$$[f(x) - g(x)]\,\varphi(x) = 0$$

следует

$$f(x) - g(x) = 0,$$

т. е.

$$f(x) = g(x).$$

Применим это к равенству (4). Если, например, $s > t$, то, сокращая обе части равенства (4) на множитель $(x-\alpha_1)^t$, мы придем к равенству, левая часть которого еще содержит множитель $x-\alpha_1$, а правая его не содержит. Выше показано, однако, что это приводит к противоречию. Таким образом, единственность разложения (2) для многочлена $f(x)$ доказана.

Объединяя вместе одинаковые множители, разложение (2) можно переписать в виде

$$f(x) = a_0 (x-\alpha_1)^{k_1} (x-\alpha_2)^{k_2} \ldots (x-\alpha_l)^{k_l}, \tag{5}$$

где

$$k_1 + k_2 + \ldots + k_l = n.$$

При этом предполагается, что среди корней $\alpha_1, \alpha_2, \ldots, \alpha_l$ уже нет равных.

Докажем, что *число k_i из* (5), $i = 1, 2, \ldots, l$, *является кратностью корня α_i в многочлене $f(x)$*. Действительно, если эта кратность равна s_i, то $k_i \leqslant s_i$. Пусть, однако, $k_i < s_i$. В силу определения кратности корня для $f(x)$ существует разложение

$$f(x) = (x-\alpha_i)^{s_i}\,\varphi(x).$$

Заменяя в этом разложении множитель $\varphi(x)$ его разложением на линейные множители, мы получили бы для $f(x)$ разложение на линейные множители, заведомо отличное от разложения (2), т. е. пришли бы к противоречию с доказанной выше единственностью этого разложения.

Мы доказали, таким образом, следующий важный результат:

Всякий многочлен $f(x)$ степени n, $n \geqslant 1$, с любыми числовыми коэффициентами имеет n корней, если каждый из корней считать столько раз, какова его кратность.

Заметим, что наша теорема справедлива и при $n = 0$, так как многочлен нулевой степени не имеет, понятно, корней. Эта теорема неприменима лишь к многочлену 0, не имеющему степени и равному нулю при любом значении x. Этим последним замечанием мы воспользуемся при доказательстве следующей теоремы:

Если многочлены $f(x)$ и $g(x)$, степени которых не превосходят n, имеют равные значения более чем при n различных значениях неизвестного, то $f(x)=g(x)$.

Действительно, многочлен $f(x)-g(x)$ имеет при наших предположениях более чем n корней, а так как его степень не превосходит n, то должно иметь место равенство $f(x)-g(x)=0$.

Таким образом, учитывая, что различных чисел бесконечно много, можно утверждать, что *для любых двух различных многочленов $f(x)$ и $g(x)$ найдутся такие значения c неизвестного x, что $f(c)\neq g(c)$.* Такие c можно найти не только среди комплексных чисел, но и среди действительных, среди рациональных и даже среди целых чисел.

Таким образом, два многочлена с числовыми коэффициентами, имеющие хотя бы при одной степени неизвестного x различные коэффициенты, будут различными комплексными функциями комплексного переменного x. Этим доказана, наконец, *равносильность для многочленов с числовыми коэффициентами двух указанных в § 20 определений равенства многочленов — алгебраического и теоретико-функционального.*

Теорема, доказанная выше, позволяет утверждать, что *многочлен, степень которого не больше n, вполне определяется своими значениями при любых различных значениях неизвестного, число которых больше n.* Можно ли эти значения многочлена задавать произвольно? Если предположить, что задаются значения многочлена при $n+1$ различных значениях неизвестного, то ответ будет положительным: *всегда существует многочлен не более чем n-й степени, принимающий наперед заданные значения при $n+1$ заданных различных значениях неизвестного.*

В самом деле, пусть нужно построить многочлен не более чем n-й степени, который при значениях неизвестного $a_1, a_2, \ldots, a_{n+1}$, предполагаемых различными, принимает соответственно значения $c_1, c_2, \ldots, c_{n+1}$. Этим многочленом будет:

$$f(x)=\sum_{i=1}^{n+1}\frac{c_i(x-a_1)\ldots(x-a_{i-1})(x-a_{i+1})\ldots(x-a_{n+1})}{(a_i-a_1)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_{n+1})}. \qquad (6)$$

Действительно, его степень не больше n, а значение $f(a_i)$ равно c_i.

Формула (6) называется *интерполяционной формулой Лагранжа.* Название «интерполяционная» связано с тем, что по этой формуле, зная значения многочлена в $n+1$ точке, можно вычислять его значения во всех других точках.

Формулы Вьета. Пусть дан многочлен $f(x)$ степени n со старшим коэффициентом 1,

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n, \qquad (7)$$

и пусть $\alpha_1, \alpha_2, \ldots, \alpha_n$ — его корни [1]. Тогда $f(x)$ обладает следующим разложением:

$$f(x) = (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n).$$

Перемножая скобки, стоящие справа, а затем приводя подобные члены и сравнивая полученные коэффициенты с коэффициентами из (7), мы получим следующие равенства, называемые *формулами Вьета* и выражающие коэффициенты многочлена через его корни:

$$\begin{aligned}
a_1 &= -(\alpha_1+\alpha_2+\ldots+\alpha_n),\\
a_2 &= \alpha_1\alpha_2+\alpha_1\alpha_3+\ldots+\alpha_1\alpha_n+\alpha_2\alpha_3+\ldots+\alpha_{n-1}\alpha_n,\\
a_3 &= -(\alpha_1\alpha_2\alpha_3+\alpha_1\alpha_2\alpha_4+\ldots+\alpha_{n-2}\alpha_{n-1}\alpha_n),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{n-1} &= (-1)^{n-1}(\alpha_1\alpha_2\ldots\alpha_{n-1}+\alpha_1\alpha_2\ldots\alpha_{n-2}\alpha_n+\ldots+\alpha_2\alpha_3\ldots\alpha_n),\\
a_n &= (-1)^n\alpha_1\alpha_2\ldots\alpha_n.
\end{aligned}$$

Таким образом, в правой части k-го равенства, $k=1, 2, \ldots, n$, стоит сумма всевозможных произведений по k корней, взятая со знаком плюс или минус, в зависимости от четности или нечетности k.

При $n=2$ эти формулы превращаются в известную из элементарной алгебры связь между корнями и коэффициентами квадратного многочлена. При $n=3$, т. е. для кубичного многочлена, эти формулы принимают вид

$$a_1 = -(\alpha_1+\alpha_2+\alpha_3), \quad a_2=\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3, \quad a_3=-\alpha_1\alpha_2\alpha_3.$$

Формулы Вьета облегчают написание многочлена по заданным его корням. Так, найдем многочлен $f(x)$ четвертой степени, имеющий простыми корнями числа 5 и -2 и двукратным корнем число 3. Мы получим:

$$\begin{aligned}
a_1 &= -(5-2+3+3) = -9,\\
a_2 &= 5\cdot(-2)+5\cdot3+5\cdot3+(-2)\cdot3+(-2)\cdot3+3\cdot3 = 17,\\
a_3 &= -[5\cdot(-2)\cdot3+5\cdot(-2)\cdot3+5\cdot3\cdot3+(-2)\cdot3\cdot3] = 33,\\
a_4 &= 5\cdot(-2)\cdot3\cdot3 = -90,
\end{aligned}$$

а поэтому

$$f(x) = x^4-9x^3+17x^2+33x-90.$$

Если старший коэффициент a_0 многочлена $f(x)$ отличен от 1, то для применения формул Вьета необходимо сначала разделить все коэффициенты на a_0, что не влияет на корни многочлена. Таким образом, в этом случае формулы Вьета дают выражение для отношений всех коэффициентов к старшему.

Многочлены с действительными коэффициентами. Сейчас будут выведены некоторые следствия из основной теоремы алгебры комплексных чисел, относящиеся к многочленам с действительными

[1] Каждый кратный корень взят здесь соответствующее число раз.

коэффициентами. По существу, именно на этих следствиях основано то исключительно большое значение основной теоремы, о котором говорилось раньше.

Пусть многочлен с действительными коэффициентами

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

имеет комплексный корень α, т. е.

$$a_0\alpha^n+a_1\alpha^{n-1}+\ldots+a_{n-1}\alpha+a_n=0.$$

Мы знаем, что последнее равенство не нарушится, если в нем все числа заменить на сопряженные. Однако все коэффициенты $a_0, a_1, \ldots$ $\ldots, a_{n-1}, a_n$, а также число 0, стоящее справа, будучи действительными, останутся при этой замене без изменения, и мы приходим к равенству

$$a_0\overline{\alpha}^n+a_1\overline{\alpha}^{n-1}+\ldots+a_{n-1}\overline{\alpha}+a_n=0,$$

т. е.

$$f(\overline{\alpha})=0.$$

Таким образом, *если комплексное (но не действительное) число α служит корнем многочлена $f(x)$ с действительными коэффициентами, то корнем для $f(x)$ будет и сопряженное число $\overline{\alpha}$.*

Многочлен $f(x)$ будет делиться, следовательно, на квадратный трехчлен

$$\varphi(x)=(x-\alpha)(x-\overline{\alpha})=x^2-(\alpha+\overline{\alpha})x+\alpha\overline{\alpha}, \tag{8}$$

коэффициенты которого, как мы знаем из § 18, действительны. Пользуясь этим, докажем, что *корни α и $\overline{\alpha}$ имеют в многочлене $f(x)$ одну и ту же кратность.*

Пусть, в самом деле, эти корни имеют соответственно кратности k и l и пусть, например, $k>l$. Тогда $f(x)$ делится на l-ю степень многочлена $\varphi(x)$,

$$f(x)=\varphi^l(x)\,q(x).$$

Многочлен $q(x)$, как частное двух многочленов с действительными коэффициентами, также имеет действительные коэффициенты, но, в противоречие с доказанным выше, он имеет число α своим $(k-l)$-кратным корнем, тогда как число $\overline{\alpha}$ не является для него корнем. Отсюда следует, что $k=l$.

Таким образом, теперь можно сказать, что *комплексные корни всякого многочлена с действительными коэффициентами попарно сопряжены.* Отсюда и из доказанной выше единственности разложений вида (2) вытекает следующий окончательный результат:

Всякий многочлен $f(x)$ с действительными коэффициентами представим, притом единственным способом (с точностью до порядка множителей), в виде произведения своего старшего коэффициента a_0 и нескольких многочленов с действительными коэф-

фициентами, линейных вида $x-\alpha$, соответствующих его действительным корням, и квадратных вида* (8), *соответствующих парам сопряженных комплексных корней.

Для дальнейшего полезно подчеркнуть, что среди многочленов с действительными коэффициентами и со старшим коэффициентом 1, неразложимыми на множители меньшей степени или, как мы будем говорить, *неприводимыми*, являются лишь линейные многочлены вида $x-\alpha$ и квадратные многочлены вида (8).

§ 25*. Рациональные дроби

В курсе математического анализа изучаются, помимо целых рациональных функций, названных нами многочленами, также *дробно-рациональные функции*; это будут частные $\frac{f(x)}{g(x)}$ двух целых рациональных функций, где $g(x)\neq 0$. Над этими функциями производятся алгебраические операции по таким же законам, как над рациональными числами, т. е. как над дробями с целыми числителями и знаменателями. Равенство двух дробно-рациональных функций или, как мы будем дальше говорить, *рациональных дробей* также понимается в том же смысле, что и равенство дробей в элементарной арифметике. Для определенности мы будем рассматривать рациональные дроби с действительными коэффициентами; читатель без труда заметит, что все содержание настоящего параграфа может быть почти дословно перенесено на случай рациональных дробей с комплексными коэффициентами.

Рациональная дробь называется *несократимой*, если ее числитель взаимно прост со знаменателем.

Всякая рациональная дробь равна некоторой несократимой дроби, определяемой однозначно с точностью до множителя нулевой степени, общего для числителя и знаменателя.

Действительно, всякую рациональную дробь можно сократить на наибольший общий делитель ее числителя и знаменателя, после чего будет получена равная ей несократимая дробь. Если, далее, равны друг другу несократимые дроби $\frac{f(x)}{g(x)}$ и $\frac{\varphi(x)}{\psi(x)}$, т. е.

$$f(x)\,\psi(x)=g(x)\,\varphi(x), \tag{1}$$

то из взаимной простоты $f(x)$ и $g(x)$ следует, по свойству б) из § 21, что $\varphi(x)$ делится на $f(x)$, а из взаимной простоты $\varphi(x)$ и $\psi(x)$ следует, что $f(x)$ делится на $\varphi(x)$. Таким образом, $f(x)=c\varphi(x)$, а тогда из (1) следует $g(x)=c\psi(x)$.

Рациональная дробь называется *правильной*, если степень числителя меньше степени знаменателя. Если к числу правильных дробей мы условимся причислять многочлен 0, то справедлива следующая теорема:

Всякая рациональная дробь представима, притом единственным способом, в виде суммы многочлена и правильной дроби.

Действительно, если дана рациональная дробь $\frac{f(x)}{g(x)}$ и если, деля числитель на знаменатель, мы получим равенство

$$f(x)=g(x)\,q(x)+r(x),$$

где степень $r(x)$ меньше степени $g(x)$, то, как легко проверить,

$$\frac{f(x)}{g(x)}=q(x)+\frac{r(x)}{g(x)}.$$

Если имеет место также равенство

$$\frac{f(x)}{g(x)}=\overline{q}(x)+\frac{\varphi(x)}{\psi(x)},$$

где степень $\varphi(x)$ меньше степени $\psi(x)$, то мы получаем равенство

$$q(x)-\overline{q}(x)=\frac{\varphi(x)}{\psi(x)}-\frac{r(x)}{g(x)}=\frac{\varphi(x)\,g(x)-\psi(x)\,r(x)}{\psi(x)\,g(x)}.$$

Так как слева стоит многочлен, а справа, как легко видеть, правильная дробь, то мы получим $q(x)-\overline{q}(x)=0$ и

$$\frac{\varphi(x)}{\psi(x)}-\frac{r(x)}{g(x)}=0.$$

Правильные рациональные дроби могут быть подвергнуты дальнейшему изучению. При этом напомним, что, как отмечено в конце предшествующего параграфа, неприводимыми действительными многочленами являются многочлены вида $x-\alpha$, где число α действительное, и многочлены вида $x^2-(\beta+\overline{\beta})\,x+\beta\overline{\beta}$, где β и $\overline{\beta}$ — пара сопряженных комплексных чисел. Как легко проверить, в комплексном случае аналогичную роль играют многочлены вида $x-\alpha$, где α — любое комплексное число.

Правильная рациональная дробь $\frac{f(x)}{g(x)}$ называется *простейшей*, если ее знаменатель $g(x)$ является степенью неприводимого многочлена $p(x)$,

$$g(x)=p^k(x), \quad k\geqslant 1,$$

а степень числителя $f(x)$ меньше степени $p(x)$.

Справедлива следующая основная теорема:

Всякая правильная рациональная дробь разлагается в сумму простейших дробей.

Доказательство. Рассмотрим сначала правильную рациональную дробь $\frac{f(x)}{g(x)\,h(x)}$, где многочлены $g(x)$ и $h(x)$ взаимно просты,

$$(g(x),\ h(x))=1.$$

Существуют, следовательно, ввиду § 21, такие многочлены $\bar{u}(x)$ и $\bar{v}(x)$, что

$$g(x)\bar{u}(x)+h(x)\bar{v}(x)=1.$$

Отсюда

$$g(x)[\bar{u}(x)f(x)]+h(x)[\bar{v}(x)f(x)]=f(x). \qquad (2)$$

Пусть, деля произведение $\bar{u}(x)f(x)$ на $h(x)$, мы получим остаток $u(x)$, степень которого меньше степени $h(x)$. Тогда равенство (2) можно будет переписать в виде

$$g(x)u(x)+h(x)v(x)=f(x), \qquad (3)$$

где $v(x)$ — многочлен, выражение которого могло бы быть без труда написано. Так как степень произведения $g(x)u(x)$ меньше степени произведения $g(x)h(x)$ и это же, по условию, верно для многочлена $f(x)$, то и произведение $h(x)v(x)$ имеет степень меньшую, чем $g(x)h(x)$, а поэтому степень $v(x)$ меньше степени $g(x)$. Из (3) вытекает теперь равенство

$$\frac{f(x)}{g(x)\,h(x)}=\frac{v(x)}{g(x)}+\frac{u(x)}{h(x)},$$

в правой части которого стоит сумма правильных дробей.

Если хотя бы один из знаменателей $g(x)$, $h(x)$ разлагается в произведение взаимно простых множителей, то можно выполнить дальнейшее разложение. Продолжая так далее, мы получим, что *всякая правильная дробь разлагается в сумму нескольких правильных дробей, каждая из которых имеет знаменателем степень некоторого неприводимого многочлена.* Точнее, если дана правильная дробь $\frac{f(x)}{g(x)}$, знаменатель которой имеет разложение на неприводимые множители

$$g(x)=p_1^{k_1}(x)p_2^{k_2}(x)\ldots p_l^{k_l}(x)$$

(всегда можно считать, конечно, что старший коэффициент знаменателя рациональной дроби равен единице), причем $p_i(x)\neq p_j(x)$ при $i\neq j$, то

$$\frac{f(x)}{g(x)}=\frac{u_1(x)}{p_1^{k_1}(x)}+\frac{u_2(x)}{p_2^{k_2}(x)}+\ldots+\frac{u_l(x)}{p_l^{k_l}(x)};$$

все слагаемые в правой части этого равенства являются правильными дробями.

Нам остается рассмотреть правильную дробь вида $\frac{u(x)}{p^k(x)}$, где $p(x)$ — неприводимый многочлен. Применяя алгоритм деления с остатком, разделим $u(x)$ на $p^{k-1}(x)$, полученный остаток разделим на $p^{k-2}(x)$ и т. д.

Мы придем к следующим равенствам:

$$u(x)=p^{k-1}(x)\,s_1(x)+u_1(x),$$
$$u_1(x)=p^{k-2}(x)\,s_2(x)+u_2(x),$$
$$\dots\dots\dots\dots\dots$$
$$u_{k-2}(x)=p(x)\,s_{k-1}(x)+u_{k-1}(x).$$

Так как при этом степень $u(x)$, по условию, меньше степени $p^k(x)$, а степень каждого из остатков $u_i(x)$, $i=1, 2, \dots, k-1$, меньше степени соответствующего делителя $p^{k-i}(x)$, то степени всех частных $s_1(x)$, $s_2(x)$, ..., $s_{k-1}(x)$ будут строго меньше степени многочлена $p(x)$. Степень последнего остатка $u_{k-1}(x)$ также меньше степени $p(x)$. Из полученных равенств следует:

$$u(x)=p^{k-1}(x)\,s_1(x)+p^{k-2}(x)\,s_2(x)+\dots+p(x)\,s_{k-1}(x)+u_{k-1}(x).$$

Отсюда мы приходим к искомому представлению рациональной дроби $\frac{u(x)}{p^k(x)}$ в виде суммы простейших дробей:

$$\frac{u(x)}{p^k(x)}=\frac{u_{k-1}(x)}{p^k(x)}+\frac{s_{k-1}(x)}{p^{k-1}(x)}+\dots+\frac{s_2(x)}{p^2(x)}+\frac{s_1(x)}{p(x)}.$$

Основная теорема доказана. Ее можно дополнить следующей теоремой единственности:

Всякая правильная рациональная дробь обладает единственным разложением в сумму простейших дробей.

Пусть, в самом деле, некоторая правильная дробь может быть двумя способами представлена в виде суммы простейших дробей. Вычитая одно из этих представлений из другого и приводя подобные члены, мы получим сумму простейших дробей, тождественно равную нулю. Пусть знаменатели простейших дробей, составляющих эту сумму, будут некоторыми степенями различных неприводимых многочленов $p_1(x)$, $p_2(x)$, ..., $p_s(x)$ и пусть наивысшая степень многочлена $p_i(x)$, $i=1, 2, \dots, s$, являющаяся одним из этих знаменателей, будет $p_i^{k_i}(x)$. Умножим обе части рассматриваемого равенства на произведение $p_1^{k_1-1}(x)\,p_2^{k_2}(x)\dots p_s^{k_s}(x)$. Все слагаемые нашей суммы, кроме одного, превратятся при этом в многочлены. Что же касается слагаемого $\frac{u(x)}{p_1^{k_1}(x)}$, то оно превратится в дробь, знаменателем которой служит $p_1(x)$, а числителем — произведение $u(x)\,p_2^{k_2}(x)\dots p_s^{k_s}(x)$. Числитель не делится нацело на знаменатель, так как многочлен $p_1(x)$ неприводим, а все множители числителя с ним взаимно просты. Выполняя деление с остатком, мы в результате получим, что равна нулю сумма многочлена и отличной от нуля правильной дроби, что, однако, невозможно.

Пример. Разложить в сумму простейших дробей действительную правильную дробь $\frac{f(x)}{g(x)}$, где

$$f(x)=2x^4-10x^3+7x^2+4x+3,$$
$$g(x)=x^5-2x^3+2x^2-3x+2.$$

Легко проверяется, что

$$g(x)=(x+2)(x-1)^2(x^2+1),$$

причем каждый из многочленов $x+2$, $x-1$, x^2+1 неприводим. Из изложенной выше теории вытекает, что искомое разложение должно иметь вид

$$\frac{f(x)}{g(x)}=\frac{A}{x+2}+\frac{B}{(x-1)^2}+\frac{C}{x-1}+\frac{Dx+E}{x^2+1}, \tag{4}$$

где числа A, B, C, D и E еще должны быть разысканы.

Из (4) вытекает равенство

$$f(x)=A(x-1)^2(x^2+1)+B(x+2)(x^2+1)+C(x+2)(x-1)(x^2+1)+$$
$$+Dx(x+2)(x-1)^2+E(x+2)(x-1)^2. \tag{5}$$

Приравнивая коэффициенты при одинаковых степенях неизвестного x из обеих частей равенства (5), мы получили бы систему пяти линейных уравнений относительно пяти неизвестных A, B, C, D, E, причем, как вытекает из доказанного выше, эта система обладает решением и притом единственным. Мы пойдем, однако, иным путем.

Полагая в равенстве (5) $x=-2$, мы придем к равенству $45A=135$, откуда

$$A=3. \tag{6}$$

Полагая, далее, в (5) $x=1$, мы получим $6B=6$, т. е.

$$B=1. \tag{7}$$

После этого положим в равенстве (5) последовательно $x=0$ и $x=-1$. Используя (6) и (7), мы получим уравнения

$$\left.\begin{aligned} -2C+2E&=-2,\\ -4C-4D+4E&=-8.\end{aligned}\right\} \tag{8}$$

Отсюда

$$D=1. \tag{9}$$

Положим, наконец, в равенстве (5) $x=2$. Используя (6), (7) и (9), мы придем к уравнению

$$20C+4E=-52,$$

которое вместе с первым из уравнений (8) дает

$$C=-2, \quad E=-3.$$

Таким образом,

$$\frac{f(x)}{g(x)}=\frac{3}{x+2}+\frac{1}{(x-1)^2}-\frac{2}{x-1}+\frac{x-3}{x^2+1}.$$

ГЛАВА ШЕСТАЯ

КВАДРАТИЧНЫЕ ФОРМЫ

§ 26. Приведение квадратичной формы к каноническому виду

Истоки теории квадратичных форм лежат в аналитической геометрии, а именно в теории кривых (и поверхностей) второго порядка. Известно, что уравнение центральной кривой второго порядка на плоскости, после перенесения начала прямоугольных координат в центр этой кривой, имеет вид

$$Ax^2+2Bxy+Cy^2=D. \tag{1}$$

Известно, далее, что можно совершить такой поворот осей координат на некоторый угол α, т. е. такой переход от координат x, y к координатам x', y':

$$\left.\begin{aligned} x&=x'\cos\alpha-y'\sin\alpha,\\ y&=x'\sin\alpha+y'\cos\alpha, \end{aligned}\right\} \tag{2}$$

что в новых координатах уравнение нашей кривой будет иметь «канонический» вид

$$A'x'^2+C'y'^2=D; \tag{3}$$

в этом уравнении коэффициент при произведении неизвестных $x'y'$ равен, следовательно, нулю. Преобразование координат (2) можно толковать, очевидно, как линейное преобразование неизвестных (см. § 13), притом невырожденное, так как определитель из его коэффициентов равен единице. Это преобразование применяется к левой части уравнения (1), и поэтому можно сказать, что левая часть уравнения (1) невырожденным линейным преобразованием (2) превращается в левую часть уравнения (3).

Многочисленные приложения потребовали построения аналогичной теории для случая, когда число неизвестных вместо двух равно любому n, а коэффициенты являются или действительными, или же любыми комплексными числами.

Обобщая выражение, стоящее в левой части уравнения (1), мы приходим к следующему понятию.

Квадратичной формой f от n неизвестных $x_1, x_2, \ldots, x_n$ называется сумма, каждый член которой является или квадратом одного из этих неизвестных, или произведением двух разных неизвестных. Квадратичная форма называется *действительной* или *комплексной* в зависимости от того, являются ли ее коэффициенты действительными или же могут быть любыми комплексными числами.

Считая, что в квадратичной форме f уже сделано приведение подобных членов, введем следующие обозначения для коэффициентов этой формы: коэффициент при x_i^2 обозначим через a_{ii}, а коэффициент при произведении $x_i x_j$ для $i \neq j$ — через $2a_{ij}$ (сравнить с (1)!). Так как, однако, $x_i x_j = x_j x_i$, то коэффициент при этом произведении мог бы быть обозначен и через $2a_{ji}$, т. е. введенные нами обозначения предполагают справедливость равенства

$$a_{ji} = a_{ij}. \tag{4}$$

Член $2a_{ij}x_i x_j$ можно записать теперь в виде

$$2a_{ij}x_i x_j = a_{ij}x_i x_j + a_{ji}x_j x_i,$$

а всю квадратичную форму f — в виде суммы всевозможных членов $a_{ij}x_i x_j$, где i и j уже независимо друг от друга принимают значения от 1 до n:

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j; \tag{5}$$

в частности, при $i = j$ получается член $a_{ii}x_i^2$.

Из коэффициентов a_{ij} можно составить, очевидно, квадратную матрицу $A = (a_{ij})$ порядка n; она называется *матрицей квадратичной формы* f, а ее ранг r — *рангом* этой квадратичной формы. Если, в частности, $r = n$, т. е. матрица — невырожденная, то и квадратичная форма f называется *невырожденной*. Ввиду равенства (4) элементы матрицы A, симметричные относительно главной диагонали, равны между собой, т. е. матрица A — *симметрическая*. Обратно, для любой симметрической матрицы A n-го порядка можно указать вполне определенную квадратичную форму (5) от n неизвестных, имеющую элементы матрицы A своими коэффициентами.

Квадратичную форму (5) можно записать в ином виде, используя введенное в § 14 умножение прямоугольных матриц. Условимся сначала о следующем обозначении: если дана квадратная или вообще прямоугольная матрица A, то через A' будет обозначаться матрица, полученная из матрицы A транспонированием. Если матрицы A и B таковы, что их произведение определено, то имеет место равенство:

$$(AB)' = B'A', \tag{6}$$

т. е. *матрица, полученная транспонированием произведения, равна произведению матриц, получающихся транспонированием сомножителей, притом взятых в обратном порядке.*

В самом деле, если произведение AB определено, то будет определено, как легко проверить, и произведение $B'A'$: число столбцов матрицы B' равно числу строк матрицы A'. Элемент матрицы $(AB)'$, стоящий в ее i-й строке и j-м столбце, в матрице AB расположен в j-й строке и i-м столбце. Он равен поэтому сумме произведений соответственных элементов j-й строки матрицы A и i-го столбца матрицы B, т. е. равен сумме произведений соответственных элементов j-го столбца матрицы A' и i-й строки матрицы B'. Этим равенство (6) доказано.

Заметим, что *матрица A тогда и только тогда будет симметрической, если она совпадает со своей транспонированной*, т. е. если

$$A' = A.$$

Обозначим теперь через X столбец, составленный из неизвестных,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

X является матрицей, имеющей n строк и один столбец. Транспонируя эту матрицу, получим матрицу

$$X' = (x_1, x_2, \ldots, x_n),$$

составленную из одной строки.

Квадратичная форма (5) с матрицей $A = (a_{ij})$ может быть записана теперь в виде следующего произведения:

$$f = X'AX. \tag{7}$$

Действительно, произведение AX будет матрицей, состоящей из одного столбца:

$$AX = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix}.$$

Умножая эту матрицу слева на матрицу X', мы получим «матрицу», состоящую из одной строки и одного столбца, а именно правую часть равенства (5).

Что произойдет с квадратичной формой f, если входящие в нее неизвестные $x_1, x_2, \ldots, x_n$ будут подвергнуты линейному преобразованию

$$x_i = \sum_{k=1}^{n} q_{ik} y_k, \quad i = 1, 2, \ldots, n, \tag{8}$$

с матрицей $Q = (q_{ik})$? Будем считать при этом, что если форма f действительная, то и элементы матрицы Q должны быть действительными. Обозначая через Y столбец из неизвестных $y_1, y_2, \ldots, y_n$, запишем линейное преобразование (8) в виде матричного равенства:

$$X = QY. \tag{9}$$

Отсюда по (6)

$$X' = Y'Q'. \tag{10}$$

Подставляя (9) и (10) в запись (7) формы f, получаем:

$$f = Y'(Q'AQ)Y,$$

или

$$f = Y'BY,$$

где

$$B = Q'AQ.$$

Матрица B будет симметрической, так как ввиду равенства (6), справедливого, очевидно, для любого числа множителей, и равенства $A' = A$, равносильного симметричности матрицы A, имеем:

$$B' = Q'A'Q = Q'AQ = B.$$

Таким образом, доказана следующая теорема:

Квадратичная форма от n неизвестных, имеющая матрицу A, после выполнения линейного преобразования неизвестных с матрицей Q превращается в квадратичную форму от новых неизвестных, причем матрицей этой формы служит произведение $Q'AQ$.

Предположим теперь, что мы выполняем невырожденное линейное преобразование, т. е. Q, а поэтому и Q' — матрицы невырожденные. Произведение $Q'AQ$ получается в этом случае умножением матрицы A на невырожденные матрицы и поэтому, как следует из результатов § 14, ранг этого произведения равен рангу матрицы A. Таким образом, *ранг квадратичной формы не меняется при выполнении невырожденного линейного преобразования.*

Рассмотрим теперь, по аналогии с указанной в начале параграфа геометрической задачей приведения уравнения центральной кривой второго порядка к каноническому виду (3), вопрос о приведении произвольной квадратичной формы некоторым невырожденным линейным преобразованием к виду суммы квадратов неизвестных, т. е. к такому виду, когда все коэффициенты при произведениях различных неизвестных равны нулю; этот специальный вид квадратичной формы называется *каноническим*. Предположим сначала, что

квадратичная форма f от n неизвестных $x_1, x_2, \ldots, x_n$ уже приведена невырожденным линейным преобразованием к каноническому виду

$$f = b_1y_1^2 + b_2y_2^2 + \ldots + b_ny_n^2, \tag{11}$$

где $y_1, y_2, \ldots, y_n$ — новые неизвестные. Некоторые из коэффициентов $b_1, b_2, \ldots, b_n$ могут, конечно, быть нулями. Докажем, что *число отличных от нуля коэффициентов в* (11) *непременно равно рангу r формы f.*

В самом деле, так как мы пришли к (11) при помощи невырожденного преобразования, то квадратичная форма, стоящая в правой части равенства (11), также должна быть ранга r. Однако матрица этой квадратичной формы имеет диагональный вид

$$\begin{pmatrix} b_1 & & & 0 \\ & b_2 & & \\ & & \ddots & \\ 0 & & & b_n \end{pmatrix},$$

и требование, чтобы эта матрица имела ранг r, равносильно предположению, что на ее главной диагонали стоит ровно r отличных от нуля элементов.

Перейдем к доказательству следующей основной теоремы о квадратичных формах.

Всякая квадратичная форма может быть приведена некоторым невырожденным линейным преобразованием к каноническому виду. Если при этом рассматривается действительная квадратичная форма, то все коэффициенты указанного линейного преобразования можно считать действительными.

Эта теорема верна для случая квадратичных форм от одного неизвестного, так как всякая такая форма имеет вид ax^2, являющийся каноническим. Мы можем, следовательно, вести доказательство индукцией по числу неизвестных, т. е. доказывать теорему для квадратичных форм от n неизвестных, считая ее уже доказанной для форм с меньшим числом неизвестных.

Пусть дана квадратичная форма

$$f = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \tag{12}$$

от n неизвестных $x_1, x_2, \ldots, x_n$. Мы постараемся найти такое невырожденное линейное преобразование, которое выделило бы из f квадрат одного из неизвестных, т. е. привело бы f к виду суммы этого квадрата и некоторой квадратичной формы от остальных неизвестных. Эта цель легко достигается в том случае, если среди коэффициентов $a_{11}, a_{22}, \ldots, a_{nn}$, стоящих в матрице формы f на главной диагонали, есть отличные от нуля, т. е. если в (12) входит

с отличным от нуля коэффициентом квадрат хотя бы одного из неизвестных x_i.

Пусть, например, $a_{11} \neq 0$. Тогда, как легко проверить, выражение $a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2$, являющееся квадратичной формой, содержит такие же члены с неизвестным x_1, как и наша форма f, а поэтому разность

$$f - a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2 = g$$

будет квадратичной формой, содержащей лишь неизвестные $x_2, \ldots, x_n$, но не x_1. Отсюда

$$f = a_{11}^{-1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2 + g.$$

Если мы введем обозначения

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \quad y_i = x_i \quad \text{при} \quad i = 2, 3, \ldots, n, \tag{13}$$

то получим

$$f = a_{11}^{-1}y_1^2 + g, \tag{14}$$

где g будет теперь квадратичной формой от неизвестных $y_2, y_3, \ldots, y_n$. Выражение (14) есть искомое выражение для формы f, так как оно получено из (12) невырожденным линейным преобразованием, а именно преобразованием, обратным линейному преобразованию (13), которое имеет своим определителем a_{11} и поэтому не вырождено.

Если же имеют место равенства $a_{11} = a_{22} = \ldots = a_{nn} = 0$, то предварительно нужно совершить вспомогательное линейное преобразование, приводящее к появлению в нашей форме f квадратов неизвестных. Так как среди коэффициентов в записи (12) этой формы должны быть отличные от нуля,— иначе нечего было бы доказывать,— то пусть, например, $a_{12} \neq 0$, т. е. f является суммой члена $2a_{12}x_1x_2$ и членов, в каждый из которых входит хотя бы одно из неизвестных $x_3, \ldots, x_n$.

Совершим теперь линейное преобразование

$$x_1 = z_1 - z_2, \quad x_2 = z_1 + z_2, \quad x_i = z_i \quad \text{при} \quad i = 3, \ldots, n. \tag{15}$$

Оно будет невырожденным, так как имеет определитель

$$\begin{vmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 2 \neq 0.$$

В результате этого преобразования член $2a_{12}x_1x_2$ нашей формы примет вид

$$2a_{12}x_1x_2 = 2a_{12}(z_1 - z_2)(z_1 + z_2) = 2a_{12}z_1^2 - 2a_{12}z_2^2,$$

т. е. в форме f появятся, с отличными от нуля коэффициентами, квадраты сразу двух неизвестных, причем они не могут сократиться ни с одним из остальных членов, так как в каждый из этих последних входит хотя бы одно из неизвестных $z_3, \ldots, z_n$. Теперь мы находимся в условиях уже рассмотренного выше случая, т. е. еще одним невырожденным линейным преобразованием можем привести форму f к виду (14).

Для окончания доказательства остается отметить, что квадратичная форма g зависит от меньшего, чем n, числа неизвестных и поэтому, по предположению индукции, некоторым невырожденным преобразованием неизвестных $y_2, y_3, \ldots, y_n$ приводится к каноническому виду. Это преобразование, рассматриваемое как (невырожденное, как легко видеть) преобразование всех n неизвестных, при котором y_1 остается без изменения, приводит, следовательно, (14) к каноническому виду. Таким образом, квадратичная форма f двумя или тремя невырожденными линейными преобразованиями, которые можно заменить одним невырожденным преобразованием — их произведением, приводится к виду суммы квадратов неизвестных с некоторыми коэффициентами. Число этих квадратов равно, как мы знаем, рангу формы r. Если, сверх того, квадратичная форма f действительная, то коэффициенты как в каноническом виде формы f, так и в линейном преобразовании, приводящем f к этому виду, будут действительными; в самом деле, и линейное преобразование, обратное (13), и линейное преобразование (15) имеют действительные коэффициенты.

Доказательство основной теоремы закончено. Метод, использованный в этом доказательстве, может быть применен в конкретных примерах для действительного приведения квадратичной формы к каноническому виду. Нужно лишь вместо индукции, которую мы использовали в доказательстве, последовательно выделять изложенным выше методом квадраты неизвестных.

Пример. Привести к каноническому виду квадратичную форму

$$f = 2x_1x_2 - 6x_2x_3 + 2x_3x_1. \tag{16}$$

Ввиду отсутствия в этой форме квадратов неизвестных мы выполним сначала невырожденное линейное преобразование

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3$$

с матрицей

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

после чего получим:

$$f = 2y_1^2 - 2y_2^2 - 4y_1y_3 - 8y_2y_3.$$

Теперь коэффициент при y_1^2 отличен от нуля, и поэтому из нашей формы можно выделить квадрат одного неизвестного. Полагая

$$z_1 = 2y_1 - 2y_3, \quad z_2 = y_2, \quad z_3 = y_3,$$

т. е. совершая линейное преобразование, для которого обратное будет иметь матрицу

$$B=\begin{pmatrix} \frac{1}{2} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

мы приведем f к виду

$$f=\frac{1}{2}z_1^2-2z_2^2-2z_3^2-8z_2z_3.$$

Пока выделился лишь квадрат неизвестного z_1, так как форма еще содержит произведение двух других неизвестных. Используя неравенство нулю коэффициента при z_2^2, еще раз применим изложенный выше метод. Совершая линейное преобразование

$$t_1=z_1,\quad t_2=-2z_2-4z_3,\quad t_3=z_3,$$

для которого обратное имеет матрицу

$$C=\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix},$$

мы приведем, наконец, форму f к каноническому виду

$$f=\frac{1}{2}t_1^2-\frac{1}{2}t_2^2+6t_3^2. \tag{17}$$

Линейное преобразование, приводящее (16) сразу к виду (17), будет иметь своей матрицей произведение

$$ABC=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 3 \\ \frac{1}{2} & -\frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Можно и непосредственной подстановкой проверить, что невырожденное $\left(\text{так как определитель равен } -\frac{1}{2}\right)$ линейное преобразование

$$\begin{aligned} x_1 &= \frac{1}{2}t_1+\frac{1}{2}t_2+3t_3, \\ x_2 &= \frac{1}{2}t_1-\frac{1}{2}t_2-t_3, \\ x_3 &= t_3 \end{aligned}$$

превращает (16) в (17).

Теория приведения квадратичной формы к каноническому виду построена по аналогии с геометрической теорией центральных кривых второго порядка, но не может считаться обобщением этой последней теории. В самом деле, в нашей теории допускается использование

любых невырожденных линейных преобразований, в то время как приведение кривой второго порядка к каноническому виду достигается применением линейных преобразований весьма специального вида (2), являющихся вращениями плоскости. Эта геометрическая теория может быть, однако, обобщена на случай квадратичных форм от n неизвестных с действительными коэффициентами. Изложение этого обобщения, называемого приведением квадратичных форм к главным осям, будет дано в гл. 8.

§ 27. Закон инерции

Канонический вид, к которому приводится данная квадратичная форма, вовсе не является для нее однозначно определенным: всякая квадратичная форма может быть приведена к каноническому виду многими различными способами. Так, рассмотренная в предшествующем параграфе квадратичная форма $f=2x_1x_2-6x_2x_3+2x_3x_1$ невырожденным линейным преобразованием

$$\begin{aligned} x_1 &= t_1+3t_2+2t_3, \\ x_2 &= t_1-t_2-2t_3, \\ x_3 &= \qquad t_2 \end{aligned}$$

приводится к каноническому виду

$$f=2t_1^2+6t_2^2-8t_3^2,$$

отличному от полученного ранее.

Возникает вопрос, что общего у тех различных канонических квадратичных форм, к которым приводится данная форма f? Этот вопрос тесно связан, как мы увидим, с таким вопросом: при каком условии одна из двух данных квадратичных форм может быть переведена в другую невырожденным линейным преобразованием? Ответ на эти вопросы зависит, однако, от того, рассматриваются ли комплексные или действительные квадратичные формы.

Предположим сначала, что рассматриваются произвольные комплексные квадратичные формы и, вместе с тем, допускается употребление невырожденных линейных преобразований также с произвольными комплексными коэффициентами. Мы знаем, что всякая квадратичная форма f от n неизвестных, имеющая ранг r, приводится к каноническому виду

$$f=c_1y_1^2+c_2y_2^2+\ldots+c_ry_r^2,$$

где все коэффициенты $c_1, c_2, \ldots, c_r$ отличны от нуля. Пользуясь тем, что из всякого комплексного числа извлекается квадратный корень, выполним следующее невырожденное линейное преобразование:

$$z_i=\sqrt{c_i}\,y_i \quad \text{при} \quad i=1, 2, \ldots, r; \; z_j=y_j \quad \text{при} \quad j=r+1, \ldots, n.$$

Оно приводит форму f к виду

$$f = z_1^2 + z_2^2 + \ldots + z_r^2, \tag{1}$$

называемому *нормальным*; это — просто сумма квадратов r неизвестных с коэффициентами, равными единице.

Нормальный вид зависит лишь от ранга r формы f, т. е. все квадратичные формы ранга r приводятся к одному и тому же нормальному виду (1). Если, следовательно, формы f и g от n неизвестных имеют одинаковый ранг r, то можно перевести f в (1), а затем (1) в g, т. е. существует невырожденное линейное преобразование, переводящее f в g. Так как, с другой стороны, никакое невырожденное линейное преобразование не изменяет ранга формы, то мы приходим к следующему результату:

Две комплексные квадратичные формы от n неизвестных тогда и только тогда переводятся друг в друга невырожденными линейными преобразованиями с комплексными коэффициентами, если эти формы имеют один и тот же ранг.

Из этой теоремы без труда вытекает, что *каноническим видом комплексной квадратичной формы ранга r может служить всякая сумма квадратов r неизвестных с любыми отличными от нуля комплексными коэффициентами.*

Положение несколько более сложно в том случае, если рассматриваются действительные квадратичные формы и, что особенно важно, допускаются лишь линейные преобразования с действительными коэффициентами. В этом случае уже не всякую форму можно привести к виду (1), так как это могло бы потребовать извлечения квадратного корня из отрицательного числа. Если, однако, мы назовем теперь *нормальным видом* квадратичной формы сумму квадратов нескольких неизвестных с коэффициентами $+1$ или -1, то легко показать, что *всякую действительную квадратичную форму f можно привести невырожденным линейным преобразованием с действительными коэффициентами к нормальному виду.*

В самом деле, форма f ранга r от n неизвестных приводится к каноническому виду, который можно записать следующим образом (меняя, если нужно, нумерацию неизвестных):

$$f = c_1 y_1^2 + \ldots + c_k y_k^2 - c_{k+1} y_{k+1}^2 - \ldots - c_r y_r^2, \qquad 0 \leqslant k \leqslant r,$$

где все числа $c_1, \ldots, c_k; c_{k+1}, \ldots, c_r$ отличны от нуля и положительны. Тогда невырожденное линейное преобразование с действительными коэффициентами

$$z_i = \sqrt{c_i}\, y_i \text{ при } i = 1, 2, \ldots, r, \quad z_j = y_j \text{ при } j = r+1, \ldots, n,$$

приводит f к нормальному виду,

$$f = z_1^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2.$$

Общее число входящих сюда квадратов будет равно рангу формы.

Действительная квадратичная форма может быть приведена к нормальному виду многими различными преобразованиями, однако с точностью до нумерации неизвестных она приводится лишь к одному нормальному виду. Это показывает следующая важная теорема, называемая *законом инерции действительных квадратичных форм*:

Число положительных и число отрицательных квадратов в нормальном виде, к которому приводится данная квадратичная форма с действительными коэффициентами действительным невырожденным линейным преобразованием, не зависят от выбора этого преобразования.

Пусть, в самом деле, квадратичная форма f ранга r от n неизвестных $x_1, x_2, \ldots, x_n$ двумя способами приведена к нормальному виду:

$$\begin{aligned} f &= y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = \\ &= z_1^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2. \end{aligned} \tag{2}$$

Так как переход от неизвестных $x_1, x_2, \ldots, x_n$ к неизвестным $y_1, y_2, \ldots, y_n$ был невырожденным линейным преобразованием, то, обратно, вторые неизвестные также будут линейно выражаться через первые с отличным от нуля определителем:

$$y_i = \sum_{s=1}^{n} a_{is} x_s, \qquad i = 1, 2, \ldots, n. \tag{3}$$

Аналогично

$$z_j = \sum_{t=1}^{n} b_{jt} x_t, \qquad j = 1, 2, \ldots, n, \tag{4}$$

причем определитель из коэффициентов снова отличен от нуля. Коэффициенты же как в (3), так и в (4) — действительные числа.

Предположим теперь, что $k < l$, и напишем систему равенств

$$y_1 = 0, \ \ldots, \ y_k = 0, \quad z_{l+1} = 0, \ \ldots, \ z_r = 0, \ \ldots, \ z_n = 0. \tag{5}$$

Если левые части этих равенств будут заменены их выражениями из (3) и (4), мы получим систему $n - l + k$ линейных однородных уравнений с n неизвестными $x_1, x_2, \ldots, x_n$. Число уравнений в этой системе меньше числа неизвестных, поэтому, как мы знаем из § 1, наша система обладает ненулевым действительным решением $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Заменим теперь в равенстве (2) все y и все z их выражениями (3) и (4), а затем подставим вместо неизвестных числа $\alpha_1, \alpha_2, \ldots, \alpha_n$. Если для краткости через $y_i(\alpha)$ и $z_j(\alpha)$ будут обозначены значения неизвестных y_i и z_j, получающиеся после такой подстановки, то (2) превращается, ввиду (5), в равенство

$$-y_{k+1}^2(\alpha) - \ldots - y_r^2(\alpha) = z_1^2(\alpha) + \ldots + z_l^2(\alpha). \tag{6}$$

Так как все коэффициенты в (3) и (4) действительные, то все квадраты, входящие в равенство (6), положительны, а поэтому (6) влечет за собой равенство нулю всех этих квадратов; отсюда следуют равенства

$$z_1(\alpha)=0, \quad \ldots, \quad z_l(\alpha)=0. \tag{7}$$

С другой стороны, по самому выбору чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$

$$z_{l+1}(\alpha)=0, \quad \ldots, \quad z_r(\alpha)=0, \quad \ldots, \quad z_n(\alpha)=0. \tag{8}$$

Таким образом, система n линейных однородных уравнений

$$z_i=0, \quad i=1, 2, \ldots, n,$$

с n неизвестными $x_1, x_2, \ldots, x_n$ обладает, ввиду (7) и (8), ненулевым решением $\alpha_1, \alpha_2, \ldots, \alpha_n$, т. е. определитель этой системы должен быть равен нулю. Это противоречит, однако, тому, что преобразование (4) предполагалось невырожденным. К такому же противоречию мы придем при $l<k$. Отсюда следует равенство $k=l$, доказывающее теорему.

Число положительных квадратов в той нормальной форме, к которой приводится данная действительная квадратичная форма f, называется *положительным индексом инерции* этой формы, число отрицательных квадратов — *отрицательным индексом инерции*, а разность между положительным и отрицательным индексами инерции — *сигнатурой* формы f. Понятно, что при заданном ранге формы задание любого из определенных сейчас трех чисел вполне определяет два других, и поэтому в дальнейших формулировках можно будет говорить о любом из этих трех чисел.

Докажем теперь следующую теорему:

Две квадратичные формы от n неизвестных с действительными коэффициентами тогда и только тогда переводятся друг в друга невырожденными действительными линейными преобразованиями, если эти формы имеют одинаковые ранги и одинаковые сигнатуры.

В самом деле, пусть форма f переводится в форму g невырожденным действительным преобразованием. Мы знаем, что это преобразование не меняет ранга формы. Оно не может менять и сигнатуры, так как в противном случае f и g приводились бы к различным нормальным видам, а тогда форма f приводилась бы, в противоречие с законом инерции, к этим обоим нормальным видам. Обратно, если формы f и g имеют одинаковые ранги и одинаковые сигнатуры, то они приводятся к одному и тому же нормальному виду и поэтому могут быть переведены друг в друга.

Если дана квадратичная форма g в каноническом виде,

$$g=b_1y_1^2+b_2y_2^2+\ldots+b_ry_r^2, \tag{9}$$

с не равными нулю действительными коэффициентами, то ранг этой формы равен, очевидно, r. Легко видеть, далее, употребляя уже применявшийся выше способ приведения такой формы к нормальному виду, что положительный индекс инерции формы g равен числу положительных коэффициентов в правой части равенства (9). Отсюда и из предшествующей теоремы вытекает такой результат:

Квадратичная форма f тогда и только тогда будет иметь форму (9) *своим каноническим видом, если ранг формы f равен r, а положительный индекс инерции этой формы совпадает с числом положительных коэффициентов в* (9).

Распадающиеся квадратичные формы. Перемножая любые две линейные формы от n неизвестных,

$$\varphi = a_1x_1 + a_2x_2 + \ldots + a_nx_n, \quad \psi = b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$

мы получим, очевидно, некоторую квадратичную форму. Не всякая квадратичная форма может быть представлена в виде произведения двух линейных форм, и мы хотим вывести условия, при которых это имеет место, т. е. при которых квадратичная форма является *распадающейся*.

Комплексная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ распадается тогда и только тогда, если ее ранг меньше или равен двум. Действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ распадается тогда и только тогда, если или ее ранг не больше единицы, или же он равен двум, а сигнатура равна нулю.

Рассмотрим сначала произведение линейных форм φ и ψ. Если хотя бы одна из этих форм нулевая, то их произведение будет квадратичной формой с нулевыми коэффициентами, т. е. оно имеет ранг 0. Если линейные формы φ и ψ пропорциональны,

$$\psi = c\varphi,$$

причем $c \neq 0$ и форма φ ненулевая, то пусть, например, коэффициент a_1 отличен от нуля. Тогда невырожденное линейное преобразование

$$y_1 = a_1x_1 + \ldots + a_nx_n, \quad y_i = x_i \quad \text{при} \quad i = 2, 3, \ldots, n$$

приводит квадратичную форму $\varphi\psi$ к виду

$$\varphi\psi = cy_1^2.$$

Справа стоит квадратичная форма ранга 1, а поэтому и квадратичная форма $\varphi\psi$ имеет ранг 1. Если же, наконец, линейные формы φ и ψ не являются пропорциональными, то пусть, например,

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

Тогда линейное преобразование

$$y_1 = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$
$$y_2 = b_1x_1 + b_2x_2 + \ldots + b_nx_n,$$
$$y_i = x_i \text{ при } i = 3, 4, \ldots, n$$

будет невырожденным; оно приводит квадратичную форму $\varphi\psi$ к виду

$$\varphi\psi = y_1y_2.$$

Справа стоит квадратичная форма ранга 2, имеющая в случае действительных коэффициентов сигнатуру 0.

Перейдем к доказательству обратного утверждения. Квадратичная форма ранга 0 может, конечно, рассматриваться как произведение двух линейных форм, одна из которых нулевая. Далее, квадратичная форма $f(x_1, x_2, \ldots, x_n)$ ранга 1 невырожденным линейным преобразованием приводится к виду

$$f = cy_1^2, \quad c \neq 0,$$

т. е. к виду

$$f = (cy_1)\,y_1.$$

Выражая y_1 линейно через $x_1, x_2, \ldots, x_n$, мы получим представление формы f в виде произведения двух линейных форм. Наконец, действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ ранга 2 и сигнатуры 0 приводится невырожденным линейным преобразованием к виду

$$f = y_1^2 - y_2^2;$$

к этому же виду может быть приведена любая комплексная квадратичная форма ранга 2. Однако

$$y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2),$$

но справа, после замены y_1 и y_2 их линейными выражениями через $x_1, x_2, \ldots, x_n$, будет стоять произведение двух линейных форм. Теорема доказана.

§ 28. Положительно определенные формы

Квадратичная форма f от n неизвестных с действительными коэффициентами называется *положительно определенной*, если она приводится к нормальному виду, состоящему из n положительных квадратов, т. е. если и ранг, и положительный индекс инерции этой формы равны числу неизвестных.

Следующая теорема дает возможность охарактеризовать положительно определенные формы, не приводя их к нормальному или каноническому виду.

Квадратичная форма f от n неизвестных $x_1, x_2, \ldots, x_n$ с действительными коэффициентами тогда и только тогда будет положительно определенной, если при всяких действительных значениях этих неизвестных, хотя бы одно из которых отлично от нуля, эта форма получает положительные значения.

Доказательство. Пусть форма f положительно определенная, т. е. приводится к нормальному виду

$$f=y_1^2+y_2^2+\ldots+y_n^2, \tag{1}$$

причем

$$y_i=\sum_{j=1}^{n} a_{ij}x_j, \quad i=1, 2, \ldots, n, \tag{2}$$

с отличным от нуля определителем из действительных коэффициентов a_{ij}. Если мы хотим подставить в f произвольные действительные значения неизвестных $x_1, x_2, \ldots, x_n$, хотя бы одно из которых отлично от нуля, то можно подставить их сначала в (2), а затем значения, полученные для всех y_i,— в (1). Заметим, что значения, полученные для $y_1, y_2, \ldots, y_n$ из (2), не могут все сразу равняться нулю, так как иначе мы получили бы, что система линейных однородных уравнений

$$\sum_{j=1}^{n} a_{ij}x_j=0, \quad i=1, 2, \ldots, n,$$

обладает ненулевым решением, хотя ее определитель отличен от нуля. Подставляя найденные для $y_1, y_2, \ldots, y_n$ значения в (1), мы получим значение формы f, равное сумме квадратов n действительных чисел, которые не все равны нулю; это значение будет, следовательно, строго положительным.

Обратно, пусть форма f не является положительно определенной, т. е. или ее ранг, или положительный индекс инерции меньше n. Это означает, что в нормальном виде этой формы, к которому она приводится, скажем, невырожденным линейным преобразованием (2), квадрат хотя бы одного из новых неизвестных, например y_n, или отсутствует совсем, или же содержится со знаком минус. Покажем, что в этом случае можно подобрать такие действительные значения для неизвестных $x_1, x_2, \ldots, x_n$, которые не все равны нулю, что значение формы f при этих значениях неизвестных равно нулю или даже отрицательно. Такими будут, например, те значения для $x_1, x_2, \ldots, x_n$, которые мы получим, решая по правилу Крамера систему линейных уравнений, получающихся из (2) при $y_1=y_2=\ldots$ $\ldots=y_{n-1}=0$, $y_n=1$. Действительно, при этих значениях неизвестных $x_1, x_2, \ldots, x_n$ форма f равна нулю, если y_n^2 не входит в нормальный вид этой формы, и равна -1, если y_n^2 входит в нормальный вид со знаком минус.

Теорема, сейчас доказанная, используется всюду, где применяются положительно определенные квадратичные формы. С ее помощью нельзя, однако, по коэффициентам формы установить, будет ли эта форма положительно определенной. Для этой цели служит другая теорема, которую мы сформулируем и докажем после того, как введем одно вспомогательное понятие.

Пусть дана квадратичная форма f от n неизвестных с матрицей $A=(a_{ij})$. Миноры порядка 1, 2, ..., n этой матрицы, расположенные в ее левом верхнем углу, т. е. миноры

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \ldots, \quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}, \quad \ldots, \quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

из которых последний совпадает, очевидно, с определителем матрицы A, называются *главными минорами* формы f.

Справедлива следующая теорема:

Квадратичная форма f от n неизвестных с действительными коэффициентами тогда и только тогда будет положительно определенной, если все ее главные миноры строго положительны.

Доказательство. При $n=1$ теорема верна, так как форма имеет в этом случае вид ax^2 и поэтому положительно определена тогда и только тогда, если $a>0$. Будем поэтому доказывать теорему для случая n неизвестных, предполагая, что для квадратичных форм от $n-1$ неизвестных она уже доказана.

Сделаем сначала следующее замечание:

Если квадратичная форма f с действительными коэффициентами, составляющими матрицу A, подвергается невырожденному линейному преобразованию с действительной матрицей Q, то *знак определителя формы (т. е. определителя ее матрицы) не меняется.*

Действительно, после преобразования мы получаем квадратичную форму с матрицей $Q'AQ$, однако, ввиду $|Q'|=|Q|$,

$$|Q'AQ|=|Q'|\cdot|A|\cdot|Q|=|A|\cdot|Q|^2,$$

т. е. определитель $|A|$ умножается на положительное число.

Пусть теперь дана квадратичная форма

$$f=\sum_{i,j=1}^{n} a_{ij}x_ix_j.$$

Ее можно записать в виде

$$f=\varphi(x_1, x_2, \ldots, x_{n-1})+2\sum_{i=1}^{n-1} a_{in}x_ix_n+a_{nn}x_n^2, \tag{3}$$

где φ будет квадратичной формой от $n-1$ неизвестных, составленной из тех членов формы f, в которые не входит неизвестное x_n.

Главные миноры формы φ совпадают, очевидно, со всеми, кроме последнего, главными минорами формы f.

Пусть форма f положительно определена. Форма φ также будет в этом случае положительно определенной: если бы существовали такие значения неизвестных $x_1, x_2, \ldots, x_{n-1}$, не все равные нулю, при которых форма φ получает не строго положительное значение, то, полагая дополнительно $x_n=0$, мы получили бы, ввиду (3), также не строго положительное значение формы f, хотя не все значения неизвестных $x_1, x_2, \ldots, x_{n-1}, x_n$ равны нулю. Поэтому, по индуктивному предположению, все главные миноры формы φ, т. е. все главные миноры формы f, кроме последнего, строго положительны. Что же касается последнего главного минора формы f, т. е. определителя самой матрицы A, то его положительность вытекает из следующих соображений: форма f, ввиду ее положительной определенности, невырожденным линейным преобразованием приводится к нормальному виду, состоящему из n положительных квадратов. Определитель этого нормального вида строго положителен, а поэтому ввиду сделанного выше замечания положителен и определитель самой формы f.

Пусть теперь строго положительны все главные миноры формы f. Отсюда вытекает положительность всех главных миноров формы φ, т. е., по индуктивному предположению, положительная определенность этой формы. Существует, следовательно, такое невырожденное линейное преобразование неизвестных $x_1, x_2, \ldots, x_{n-1}$, которое приводит форму φ к виду суммы $n-1$ положительных квадратов от новых неизвестных $y_1, y_2, \ldots, y_{n-1}$. Это линейное преобразование можно дополнить до (невырожденного) линейного преобразования всех неизвестных $x_1, x_2, \ldots, x_n$, полагая $x_n=y_n$. Ввиду (3) форма f приводится указанным преобразованием к виду

$$f=\sum_{i=1}^{n-1} y_i^2+2\sum_{i=1}^{n-1} b_{in}y_iy_n+b_{nn}y_n^2; \tag{4}$$

точные выражения коэффициентов b_{in} для нас несущественны. Так как

$$y_i^2+2b_{in}y_iy_n=(y_i+b_{in}y_n)^2-b_{in}^2y_n^2,$$

то невырожденное линейное преобразование

$$\begin{aligned} z_i&=y_i+b_{in}y_n, \quad i=1,\ 2,\ \ldots,\ n-1,\\ z_n&=\qquad y_n \end{aligned}$$

приводит, ввиду (4), форму f к каноническому виду

$$f=\sum_{i=1}^{n-1} z_i^2+cz_n^2. \tag{5}$$

Для доказательства положительной определенности формы f остается доказать положительность числа c. Определитель формы, стоящей в правой части равенства (5), равен c. Этот определитель должен, однако, быть положительным, так как правая часть равенства (5) получена из формы f двумя невырожденными линейными преобразованиями, а определитель формы f был, как последний из главных миноров этой формы, положительным.

Доказательство теоремы закончено.

Примеры 1. Квадратичная форма

$$f = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

положительно определена, так как ее главные миноры

$$5, \quad \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1$$

положительны.

2. Квадратичная форма

$$f = 3x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$$

не будет положительно определенной, так как ее второй главный минор отрицателен:

$$\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = -1.$$

Заметим, что по аналогии с положительно определенными квадратичными формами можно ввести *отрицательно определенные формы*, т. е. такие невырожденные квадратичные формы с действительными коэффициентами, нормальный вид которых содержит лишь отрицательные квадраты неизвестных. Вырожденные квадратичные формы, нормальный вид которых состоит из квадратов одного знака, называются иногда *полуопределенными*. Наконец, *неопределенными* будут такие квадратичные формы, нормальный вид которых содержит как положительные, так и отрицательные квадраты неизвестных.

ГЛАВА СЕДЬМАЯ

ЛИНЕЙНЫЕ ПРОСТРАНСТВА

§ 29. Определение линейного пространства. Изоморфизм

Определение n-мерного векторного пространства, данное в § 8, начиналось с определения n-мерного вектора как упорядоченной системы n чисел. Для n-мерных векторов были введены затем сложение и умножение на числа, что и привело к понятию n-мерного векторного пространства. Первыми примерами векторных пространств являются совокупности векторов-отрезков, выходящих из начала координат на плоскости или в трехмерном пространстве. Однако, встречаясь с этими примерами в курсе геометрии, мы не всегда считаем необходимым задавать векторы их компонентами в некоторой фиксированной системе координат, так как и сложение векторов и их умножение на скаляр определяются геометрически, независимо от выбора системы координат. Именно, сложение векторов на плоскости или в пространстве производится по правилу параллелограмма, а умножение вектора на число α означает растяжение этого вектора в α раз (с изменением направления вектора на противоположное, если α отрицательно). Целесообразно и в общем случае дать «бескоординатное» определение векторного пространства, т. е. определение, не требующее задания векторов упорядоченными системами чисел. Сейчас будет дано такое определение. Это определение является аксиоматическим: в нем ничего не будет сказано о свойствах отдельного вектора, но будут перечислены те свойства, которыми должны обладать операции над векторами.

Пусть дано множество V; его элементы будут обозначаться малыми латинскими буквами: a, b, c, ...[1]. Пусть, далее, в множестве V определены *операция сложения*, ставящая в соответствие всякой паре элементов a, b из V однозначно определенный элемент $a+b$ из V, называемый их *суммой*, и *операция умножения на действительное число*, причем *произведение* αa *элемента* a на число α однозначно определено и принадлежит к V.

---

[1]) В отличие от того, что было принято в гл. 2, мы будем в настоящей и следующей главах обозначать векторы малыми латинскими буквами, а числа — малыми греческими буквами.

Элементы множества V будут называться *векторами*, а само V — *действительным линейным* (или *векторным*, или *аффинным*) *пространством*, если указанные операции обладают следующими свойствами I—VIII:

I. Сложение коммутативно, $a+b=b+a$.

II. Сложение ассоциативно, $(a+b)+c=a+(b+c)$.

III. В V существует *нулевой элемент* 0, удовлетворяющий условию: $a+0=a$ для всех a из V.

Легко доказать, используя I, *единственность нулевого элемента:* если 0_1 и 0_2 — два нулевых элемента, то

$$0_1+0_2=0_1,$$
$$0_1+0_2=0_2+0_1=0_2,$$

откуда $0_1=0_2$.

IV. Для всякого элемента a в V существует *противоположный элемент* $-a$, удовлетворяющий условию: $a+(-a)=0$.

Легко проверяется, ввиду II и I, *единственность противоположного элемента:* если $(-a)_1$ и $(-a)_2$ — два противоположных элемента для a, то

$$(-a)_1+[a+(-a)_2]=(-a_1)+0=(-a)_1,$$
$$[(-a)_1+a]+(-a)_2=0+(-a)_2=(-a)_2,$$

откуда $(-a)_1=(-a)_2$.

Из аксиом I—IV выводится *существование и единственность разности* $a-b$, т. е. такого элемента, который удовлетворяет уравнению

$$b+x=a. \tag{1}$$

Действительно, можно положить

$$a-b=a+(-b),$$

так как

$$b+[a+(-b)]=[b+(-b)]+a=0+a=a.$$

Если же существует еще такой элемент c, который удовлетворяет уравнению (1), т. е.

$$b+c=a,$$

то, прибавляя к обеим частям этого равенства элемент $-b$, получаем, что

$$c=a+(-b).$$

Дальнейшие аксиомы V—VIII (ср. § 8) связывают умножение на число со сложением и с операциями над числами. Именно, для любых элементов a, b из V, для любых действительных чисел α, β

и для действительного числа 1 должны иметь место равенства:

V. $\alpha(a+b)=\alpha a+\alpha b$;

VI. $(\alpha+\beta)a=\alpha a+\beta a$;

VII. $(\alpha\beta)a=\alpha(\beta a)$;

VIII. $1\cdot a=a$.

Укажем некоторые простейшие следствия из этих аксиом.

[1]. $\alpha\cdot 0=0$.

Действительно, для некоторого a из V

$$\alpha a=\alpha(a+0)=\alpha a+\alpha\cdot 0,$$

т. е.

$$\alpha\cdot 0=\alpha a-\alpha a=\alpha a+[-(\alpha a)]=0.$$

[2]. $0\cdot a=0$,

где слева стоит число нуль, а справа — нулевой элемент из V.

Для доказательства возьмем любое число α. Тогда

$$\alpha a=(\alpha+0)a=\alpha a+0\cdot a,$$

откуда

$$0\cdot a=\alpha a-\alpha a=0.$$

[3]. Если $\alpha a=0$, то или $\alpha=0$, или $a=0$.

Действительно, если $\alpha\neq 0$, т. е. число α^{-1} существует, то

$$a=1\cdot a=(\alpha^{-1}\alpha)a=\alpha^{-1}(\alpha a)=\alpha^{-1}\cdot 0=0.$$

[4]. $\alpha(-a)=-\alpha a$.

В самом деле,

$$\alpha a+\alpha(-a)=\alpha[a+(-a)]=\alpha\cdot 0=0,$$

т. е. элемент $\alpha(-a)$ противоположен элементу αa.

[5]. $(-\alpha)a=-\alpha a$.

Действительно,

$$\alpha a+(-\alpha)a=[\alpha+(-\alpha)]a=0\cdot a=0,$$

т. е. элемент $(-\alpha)a$ противоположен элементу αa.

[6]. $\alpha(a-b)=\alpha a-\alpha b$.

Действительно, по [4],

$$\alpha(a-b)=\alpha[a+(-b)]=\alpha a+\alpha(-b)=\alpha a+(-\alpha b)=\alpha a-\alpha b.$$

[7]. $(\alpha-\beta)a=\alpha a-\beta a$.

В самом деле,

$$(\alpha-\beta)a=[\alpha+(-\beta)]a=\alpha a+(-\beta)a=\alpha a+(-\beta a)=\alpha a-\beta a.$$

Заметим, что перечисленными выше аксиомами и следствиями из них мы будем пользоваться дальше без специальных оговорок.

Выше дано определение действительного линейного пространства. Если бы мы предположили, что в множестве V определено умножение не только на действительные, но и на любые комплексные числа, то, сохраняя те же аксиомы I—VIII, получили бы определение *комплексного линейного пространства*. Для определенности ниже рассматриваются действительные линейные пространства, однако все, что будет сказано в настоящей главе, переносится дословно на случай комплексных линейных пространств.

Примеры действительных линейных пространств могут быть легко указаны. Ими будут, прежде всего, те n-мерные действительные векторные пространства, составленные из векторов-строк, которые изучались в гл. 2. Линейными пространствами будут и множества векторов-отрезков, выходящих из начала координат на плоскости или в трехмерном пространстве, если операции сложения и умножения на число понимать в том геометрическом смысле, который был указан в начале параграфа.

Существуют также примеры линейных пространств, так сказать «бесконечномерных». Рассмотрим всевозможные последовательности действительных чисел; они имеют вид

$$a = (\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots).$$

Операции над последовательностями будем производить покомпонентно: если

$$b = (\beta_1, \beta_2, \ldots, \beta_n, \ldots),$$

то

$$a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n, \ldots);$$

с другой стороны, для любого действительного числа γ

$$\gamma a = (\gamma\alpha_1, \gamma\alpha_2, \ldots, \gamma\alpha_n, \ldots).$$

Все аксиомы I—VIII выполняются, т. е. мы получаем действительное линейное пространство.

Примером бесконечномерного пространства будет также множество всевозможных действительных функций действительного переменного, если сложение функций и их умножение на действительное число понимать так, как это принято в теории функций, т. е. как сложение или умножение на число значений функций при каждом значении независимого переменного.

Изоморфизм. Нашей ближайшей целью будет выделение среди всех линейных пространств тех, которые естественно назвать конечномерными. Введем сначала одно общее понятие.

В определении линейного пространства говорилось о свойствах операций над векторами, но ничего не говорилось о свойствах

самих векторов. Ввиду этого может случиться, что хотя векторы некоторых двух данных линейных пространств по своей природе совершенно различны, однако с точки зрения свойств операций эти два пространства неразличимы. Точное определение таково:

Два действительных линейных пространства V и V' называются *изоморфными*, если между их векторами установлено взаимно однозначное соответствие — всякому вектору a из V сопоставлен вектор a' из V', образ вектора a, причем различные векторы из V обладают различными образами и всякий вектор из V' служит образом некоторого вектора из V, — и если при этом соответствии образом суммы двух векторов служит сумма образов этих векторов,

$$(a+b)'=a'+b', \tag{2}$$

а образом произведения вектора на число служит произведение образа этого вектора на то же число,

$$(\alpha a)'=\alpha a'. \tag{3}$$

Отметим, что взаимно однозначное соответствие между пространствами V и V', удовлетворяющее условиям (2) и (3), называется *изоморфным соответствием*.

Так, пространство векторов-отрезков на плоскости, выходящих из начала координат, изоморфно двумерному векторному пространству, составленному из упорядоченных пар действительных чисел: мы получим изоморфное соответствие между этими пространствами, если на плоскости фиксируем некоторую систему координат и всякому вектору-отрезку сопоставим упорядоченную пару его координат.

Докажем следующее свойство изоморфизма линейных пространств: *образом нуля пространства V при изоморфном соответствии между пространствами V и V' служит нуль пространства V'.*

Пусть, в самом деле, a будет некоторый вектор из V, a' — его образ в V'. Тогда, ввиду (2),

$$a'=(a+0)'=a'+0',$$

т. е. $0'$ будет нулем пространства V'.

§ 30. Конечномерные пространства. Базы

Как читатель без труда может проверить, те два определения линейной зависимости векторов-строк, которые были даны в § 9, равно как и доказательство эквивалентности этих определений, используют лишь операции над векторами и поэтому могут быть перенесены на случай любых линейных пространств. В аксиоматически определенных линейных пространствах можно говорить, следовательно, о линейно независимых системах векторов, о макси-

мальных линейно независимых системах, если такие существуют, и т. д.

Если линейные пространства V и V' изоморфны, то система векторов $a_1, a_2, \ldots, a_k$ из V тогда и только тогда линейно зависима, если линейно зависима система их образов $a'_1, a'_2, \ldots, a'_k$ в V'.

Заметим, что если соответствие $a \to a'$ (для всех a из V) является изоморфным соответствием между V и V', то и обратное соответствие $a' \to a$ также будет изоморфным. Поэтому достаточно рассмотреть случай, когда линейно зависима система $a_1, a_2, \ldots, a_k$. Пусть существуют такие числа $\alpha_1, \alpha_2, \ldots, \alpha_k$, не все равные нулю, что

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0.$$

Образом правой части этого равенства при рассматриваемом изоморфизме служит, как мы знаем, нуль $0'$ пространства V'. Беря образ левой части и применяя несколько раз (2) и (3), получаем

$$\alpha_1 a'_1 + \alpha_2 a'_2 + \ldots + \alpha_k a'_k = 0',$$

т. е. система $a'_1, a'_2, \ldots, a'_k$ также оказалась линейно зависимой.

Конечномерные пространства. Линейное пространство V называется *конечномерным*, если в нем можно найти конечную максимальную линейно независимую систему векторов; всякая такая система векторов будет называться *базой* пространства V.

Конечномерное линейное пространство может обладать многими различными базами. Так, в пространстве векторов-отрезков на плоскости базой служит любая пара векторов, отличных от нуля и не лежащих на одной прямой. Заметим, что наше определение конечномерного пространства не дает пока ответа на вопрос, могут ли в этом пространстве существовать базы, состоящие из разного числа векторов. Больше того, можно было бы допустить даже, что в некоторых конечномерных пространствах существуют базы со сколь угодно большим числом векторов. Сейчас мы приступим к выяснению того, каково же положение на самом деле.

Пусть линейное пространство V обладает базой

$$e_1, e_2, \ldots, e_n, \tag{1}$$

состоящей из n векторов. Если a — произвольный вектор из V, то из максимальности линейно независимой системы (1) следует, что a линейно выражается через эту систему,

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n. \tag{2}$$

С другой стороны, ввиду линейной независимости системы (1) выражение (2) будет для вектора a единственным: если

$$a = \alpha'_1 e_1 + \alpha'_2 e_2 + \ldots + \alpha'_n e_n,$$

то

$$(\alpha_1-\alpha_1')\, e_1+(\alpha_2-\alpha_2')\, e_2+\ldots+(\alpha_n-\alpha_n')\, e_n=0,$$

откуда

$$\alpha_i=\alpha_i', \quad i=1,\ 2,\ \ldots,\ n.$$

Таким образом, вектору a однозначно соответствует строка

$$(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n) \tag{3}$$

коэффициентов его выражения (2) через базу (1) или, как мы будем говорить, *строка его координат в базе* (1). Обратно, всякая строка вида (3), т. е. всякий n-мерный вектор в смысле гл. 2, служит строкой координат в базе (1) для некоторого вектора пространства V, а именно для вектора, записывающегося через базу (1) в виде (2).

Мы получили, следовательно, взаимно однозначное соответствие между всеми векторами пространства V и всеми векторами n-мерного векторного пространства строк. Покажем, что это соответствие, зависящее, понятно, от выбора базы (1), является изоморфным.

Возьмем в пространстве V, помимо вектора a, выражающегося через базу (1) в виде (2), также вектор b, выражение которого через базу (1) будет

$$b=\beta_1e_1+\beta_2e_2+\ldots+\beta_ne_n.$$

Тогда

$$a+b=(\alpha_1+\beta_1)\, e_1+(\alpha_2+\beta_2)\, e_2+\ldots+(\alpha_n+\beta_n)\, e_n,$$

т. е. *сумме векторов a и b соответствует сумма строк их координат в базе* (1). С другой стороны,

$$\gamma a=(\gamma\alpha_1)\, e_1+(\gamma\alpha_2)\, e_2+\ldots+(\gamma\alpha_n)\, e_n,$$

т. е. *произведению вектора a на число γ соответствует произведение строки его координат в базе* (1) *на это же число γ*.

Этим доказана следующая теорема:

Всякое линейное пространство, обладающее базой из n векторов, изоморфно n-мерному векторному пространству строк.

Как мы знаем, при изоморфном соответствии между линейными пространствами линейно зависимая система векторов переходит в линейно зависимую и обратно, а поэтому линейно независимая переходит в линейно независимую. Отсюда следует, что *при изоморфном соответствии база переходит в базу.*

В самом деле, пусть база $e_1,\ e_2,\ \ldots,\ e_n$ пространства V переходит при изоморфном соответствии между пространствами V и V' в систему векторов $e_1',\ e_2',\ \ldots,\ e_n'$ пространства V', которая хотя и линейно независима, но не является максимальной. В V' можно найти, следовательно, такой вектор f', что система $e_1',\ e_2',\ \ldots,\ e_n',$ f' остается линейно независимой. Вектор f' служит, однако, образом при рассматриваемом изоморфизме для некоторого вектора f из V.

Мы получаем, что система векторов $e_1, e_2, \ldots, e_n, f$ должна быть линейно независимой в противоречие с определением базы.

Мы знаем, далее (см. § 9), что в n-мерном векторном пространстве строк все максимальные линейно независимые системы состоят из n векторов, что всякая система из $n+1$ вектора линейно зависима и что всякая линейно независимая система векторов содержится в некоторой максимальной линейно независимой системе. Используя установленные выше свойства изоморфных соответствий, мы приходим к следующим результатам:

Все базы конечномерного линейного пространства V состоят из одного и того же числа векторов. Если это число равно n, то V будет называться *n-мерным линейным пространством*, а число n — *размерностью* этого пространства.

Всякая система из $n+1$ вектора n-мерного линейного пространства линейно зависима.

Всякая линейно независимая система векторов n-мерного линейного пространства содержится в некоторой базе этого пространства.

Теперь легко проверить, что указанные выше примеры действительных линейных пространств — пространство последовательностей и пространство функций — не являются конечномерными пространствами: в каждом из этих пространств читатель без труда найдет линейно независимые системы, состоящие из сколь угодно большого числа векторов.

Связь между базами. Объектом изучения являются для нас конечномерные линейные пространства. Понятно, что, изучая n-мерные линейные пространства, мы по существу изучаем то n-мерное векторное пространство строк, которое было введено еще в гл. 2. Однако раньше в этом пространстве была выделена одна база — а именно база, составленная из единичных векторов, т. е. векторов, у которых одна координата равна единице, а все остальные координаты равны нулю, — и все векторы пространства задавались строками их координат в этой базе; теперь же все базы пространства являются для нас равноправными.

Посмотрим, как много баз можно найти в n-мерном линейном пространстве и как эти базы связаны друг с другом.

Пусть в n-мерном линейном пространстве V заданы базы

$$e_1, \quad e_2, \ldots, e_n \tag{4}$$

и

$$e'_1, \quad e'_2, \ldots, e'_n. \tag{5}$$

Каждый вектор базы (5), как и всякий вектор пространства V, однозначно записывается через базу (4),

$$e'_i = \sum_{j=1}^{n} \tau_{ij} e_j, \quad i = 1, 2, \ldots, n. \tag{6}$$

Матрица

$$T=\begin{pmatrix}\tau_{11} & \dots & \tau_{1n}\\ \cdot & \cdot & \cdot \\ \tau_{n1} & \dots & \tau_{nn}\end{pmatrix},$$

строки которой являются строками координат векторов (5) в базе (4), называется *матрицей перехода* от базы (4) к базе (5).

Связь между базами (4) и (5) и матрицей перехода T можно записать, ввиду (6), в виде матричного равенства

$$\begin{pmatrix} e_1' \\ e_2' \\ \cdot \\ \cdot \\ \cdot \\ e_n' \end{pmatrix} = \begin{pmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \dots & \tau_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \tau_{n1} & \tau_{n2} & \dots & \tau_{nn} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{pmatrix} \tag{7}$$

или, обозначая базы (4) и (5), записанные в столбец, соответственно через e и e', в виде

$$e' = Te.$$

С другой стороны, если T' — матрица перехода от базы (5) к базе (4), то

$$e = T'e'.$$

Отсюда

$$e = (T'T)\,e,$$
$$e' = (TT')\,e',$$

т. е., ввиду линейной независимости баз e и e',

$$T'T = TT' = E,$$

откуда

$$T' = T^{-1}.$$

Этим доказано, что *матрица перехода от одной базы к другой всегда является невырожденной матрицей.*

Всякая невырожденная квадратная матрица порядка n с действительными элементами служит матрицей перехода от данной базы n-мерного действительного линейного пространства к некоторой другой базе.

Пусть, в самом деле, дана база (4) и невырожденная матрица T порядка n. Возьмем в качестве (5) систему векторов, для которых строки матрицы T служат строками координат в базе (4); имеет место, следовательно, равенство (7). Векторы (5) линейно независимы — линейная зависимость между ними влекла бы за собой линейную зависимость строк матрицы T в противоречие с ее невырожденностью. Поэтому система (5), как линейно независимая

система, состоящая из n векторов, является базой нашего пространства, а матрица T служит матрицей перехода от базы (4) к базе (5).

Мы видим, что в n-мерном линейном пространстве можно найти столь же много различных баз, как много существует различных невырожденных квадратных матриц порядка n. Правда, при этом две базы, состоящие из одних и тех же векторов, но записанных в различном порядке, считаются различными.

Преобразование координат вектора. Пусть в n-мерном линейном пространстве даны базы (4) и (5) с матрицей перехода $T=(\tau_{ij})$,

$$e' = Te.$$

Найдем связь между строками координат произвольного вектора a в этих базах.

Пусть

$$a = \sum_{j=1}^{n} \alpha_j e_j, \tag{8}$$

$$a = \sum_{i=1}^{n} \alpha'_i e'_i.$$

Используя (6), получаем:

$$a = \sum_{i=1}^{n} \alpha'_i \left(\sum_{j=1}^{n} \tau_{ij} e_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \alpha'_i \tau_{ij} \right) e_j.$$

Сравнивая с (8) и используя единственность записи вектора через базу, получаем:

$$\alpha_j = \sum_{i=1}^{n} \alpha'_i \tau_{ij}, \qquad j = 1, 2, \ldots, n,$$

т. е. имеет место матричное равенство

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha'_1, \alpha'_2, \ldots, \alpha'_n)\, T.$$

Таким образом, *строка координат вектора a в базе e равна строке координат этого вектора в базе e', умноженной справа на матрицу перехода от базы e к базе e'.*

Отсюда следует, понятно, равенство

$$(\alpha'_1, \alpha'_2, \ldots, \alpha'_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)\, T^{-1}.$$

Пример. Рассмотрим трехмерное действительное линейное пространство с базой

$$e_1, e_2, e_3. \tag{9}$$

Векторы

$$\left.\begin{aligned} e'_1 &= 5e_1 - e_2 - 2e_3, \\ e'_2 &= 2e_1 + 3e_2, \\ e'_3 &= -2e_1 + e_2 + e_3 \end{aligned}\right\} \tag{10}$$

также составляют базу в этом пространстве, причем матрицей перехода от (9) к (10) служит матрица

$$T=\begin{pmatrix} 5 & -1 & -2 \\ 2 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix},$$

откуда

$$T^{-1}=\begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}.$$

Вектор

$$a=e_1+4e_2-e_3$$

имеет поэтому в базе (10) строку координат

$$(\alpha'_1,\ \alpha'_2,\ \alpha'_3)=(1,\ 4,\ -1)\begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}=(-13,\ 6,\ -27),$$

т. е.

$$a=-13e'_1+6e'_2-27e'_3.$$

§ 31. Линейные преобразования

В гл. 3 мы уже встречались с понятием линейного преобразования неизвестных. Понятие, которое будет сейчас введено, носит такое же название, но имеет иной характер. Впрочем, некоторые связи между этими двумя одноименными понятиями могли бы быть указаны.

Пусть дано n-мерное действительное линейное пространство, которое обозначим через V_n. Рассмотрим *преобразование* этого пространства, т. е. отображение, переводящее каждый вектор a пространства V_n в некоторый вектор a' этого же пространства. Вектор a' называется *образом* вектора a при рассматриваемом преобразовании.

Если преобразование обозначено через φ, то образ вектора a условимся записывать не через $\varphi(a)$ или φa, что читателю было бы привычнее, а через $a\varphi$. Таким образом,

$$a'=a\varphi.$$

Преобразование φ линейного пространства V_n называется *линейным преобразованием* этого пространства, если сумму любых двух векторов a, b оно переводит в сумму образов этих векторов,

$$(a+b)\varphi=a\varphi+b\varphi, \qquad (1)$$

а произведение любого вектора a на любое число α переводит в произведение образа вектора a на это же число α,

$$(\alpha a)\varphi=\alpha(a\varphi). \qquad (2)$$

Из этого определения немедленно вытекает, что *линейное преобразование линейного пространства переводит любую линейную*

комбинацию данных векторов $a_1, a_2, \ldots, a_k$ *в линейную комбинацию (с теми же коэффициентами) образов этих векторов,*

$$(\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k)\varphi = \alpha_1(a_1\varphi) + \alpha_2(a_2\varphi) + \ldots + \alpha_k(a_k\varphi). \quad (3)$$

Докажем следующее утверждение:

При любом линейном преобразовании φ *линейного пространства* V_n *нулевой вектор* 0 *остается неподвижным,*

$$0\varphi = 0,$$

а образом вектора, противоположного для данного вектора a, *служит вектор, противоположный для образа вектора* a,

$$(-a)\varphi = -a\varphi.$$

В самом деле, если b — произвольный вектор, то, ввиду (2),

$$0\varphi = (0 \cdot b)\varphi = 0 \cdot (b\varphi) = 0.$$

С другой стороны,

$$(-a)\varphi = [(-1)a]\varphi = (-1)(a\varphi) = -a\varphi.$$

Понятие линейного преобразования линейного пространства возникло как обобщение известного из курса аналитической геометрии понятия аффинного преобразования плоскости или трехмерного пространства; действительно, условия (1) и (2) для аффинных преобразований выполняются. Эти условия выполняются и для проекций векторов на плоскости или в трехмерном пространстве на некоторую прямую (или на некоторую плоскость). Таким образом, например, в двумерном линейном пространстве векторов-отрезков, выходящих из начала координат плоскости, преобразование, переводящее всякий вектор в его проекцию на некоторую ось, проходящую через начало координат, будет линейным преобразованием.

Примерами линейных преобразований в произвольном пространстве V_n служат *тождественное преобразование* ε, оставляющее всякий вектор a на месте,

$$a\varepsilon = a,$$

и *нулевое преобразование* ω, отображающее всякий вектор a в нуль,

$$a\omega = 0.$$

Сейчас будет получено некоторое обозрение всех линейных преобразований линейного пространства V_n. Пусть

$$e_1, e_2, \ldots, e_n \quad (4)$$

— база этого пространства; как и раньше, базу (4), расположенную в столбец, будем обозначать через e. Так как всякий вектор a пространства V_n однозначно представляется в виде линейной комбинации

векторов базы (4), то, ввиду (3), образ вектора a с теми же коэффициентами выражается через образы векторов (4). Иными словами, *всякое линейное преобразование* φ *пространства* V_n *однозначно определяется заданием образов* $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ *всех векторов фиксированной базы* (4).

Какова бы ни была упорядоченная система из n *векторов пространства* V_n,

$$c_1, \ c_2, \ \ldots, \ c_n, \tag{5}$$

существует, притом единственное, такое линейное преобразование φ *этого пространства, что* (5) *служит системой образов векторов базы* (4) *при этом преобразовании,*

$$e_i\varphi = c_i, \quad i = 1, 2, \ldots, n. \tag{6}$$

Единственность преобразования φ уже доказана выше и нужно доказать лишь его существование. Определим преобразование φ следующим образом: если a — произвольный вектор пространства и

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

— его запись в базе (4), то положим

$$a\varphi = \sum_{i=1}^{n} \alpha_i c_i. \tag{7}$$

Докажем линейность этого преобразования. Если

$$b = \sum_{i=1}^{n} \beta_i e_i$$

— любой другой вектор пространства, то

$$(a+b)\varphi = \left[\sum_{i=1}^{n} (\alpha_i + \beta_i) e_i\right]\varphi = \sum_{i=1}^{n} (\alpha_i + \beta_i) c_i = \sum_{i=1}^{n} \alpha_i c_i + \sum_{i=1}^{n} \beta_i c_i = a\varphi + b\varphi.$$

Если же γ — любое число, то

$$(\gamma a)\varphi = \left[\sum_{i=1}^{n} (\gamma\alpha_i) e_i\right]\varphi = \sum_{i=1}^{n} (\gamma\alpha_i) c_i = \gamma \sum_{i=1}^{n} \alpha_i c_i = \gamma(a\varphi).$$

Что же касается справедливости равенств (6), то она вытекает из определения (7) преобразования φ, так как все координаты вектора e_i в базе (4) равны нулю, кроме i-й координаты, равной единице.

Нами установлено, следовательно, взаимно однозначное соответствие между всеми линейными преобразованиями линейного пространства V_n *и всеми упорядоченными системами* (5) *из* n *векторов этого пространства.*

Всякий вектор c_i обладает, однако, определенной записью в базе (4),

$$c_i = \sum_{j=1}^{n} \alpha_{ij} e_j, \quad i = 1, 2, \ldots, n. \tag{8}$$

Из координат вектора c_i в базе (4) можно составить квадратную матрицу

$$A = (\alpha_{ij}), \tag{9}$$

беря в качестве ее i-й строки строку координат вектора c_i, $i = 1, 2, \ldots$ $\ldots, n$. Так как система (5) была произвольной, то матрица A будет произвольной квадратной матрицей порядка n с действительными элементами.

Мы имеем, таким образом, взаимно однозначное соответствие между всеми линейными преобразованиями пространства V_n и всеми квадратными матрицами порядка n; это соответствие зависит, конечно, от выбора базы (4).

Будем говорить, что матрица A *задает* линейное преобразование φ в базе (4), или, короче, что A есть *матрица линейного преобразования* φ в базе (4). Если через $e\varphi$ мы обозначим столбец, составленный из образов векторов базы (4), то из (6), (8) и (9) вытекает следующее матричное равенство, полностью описывающее связи, существующие между линейным преобразованием φ, базой e и матрицей A, задающей это линейное преобразование в этой базе:

$$e\varphi = Ae. \tag{10}$$

Покажем, как, зная матрицу A линейного преобразования φ в базе (4), по координатам вектора a в этой базе найти координаты его образа $a\varphi$. Если

$$a = \sum_{i=1}^{n} \alpha_i e_i,$$

то

$$a\varphi = \sum_{i=1}^{n} \alpha_i (e_i\varphi),$$

что равносильно матричному равенству

$$a\varphi = (\alpha_1, \alpha_2, \ldots, \alpha_n)(e\varphi).$$

Используя (10) и учитывая, что ассоциативность умножения матриц легко проверяется и в том случае, когда одна из матриц является столбцом, составленным из векторов, мы получаем:

$$a\varphi = [(\alpha_1, \alpha_2, \ldots, \alpha_n) A] e.$$

Отсюда следует, что *строка координат вектора $a\varphi$ равна строке координат вектора a, умноженной справа на матрицу A линейного преобразования φ, все в базе* (4).

Пример. Пусть в базе e_1, e_2, e_3 трехмерного линейного пространства линейное преобразование φ задается матрицей

$$A=\begin{pmatrix}-2 & 1 & 0\\ 1 & 3 & 2\\ 0 & -4 & 1\end{pmatrix}.$$

Если

$$a=5e_1+e_2-2e_3,$$

то

$$(5,\ 1,\ -2)\begin{pmatrix}-2 & 1 & 0\\ 1 & 3 & 2\\ 0 & -4 & 1\end{pmatrix}=(-9,\ 16,\ 0),$$

т. е.

$$a\varphi=-9e_1+16e_2.$$

Связь между матрицами линейного преобразования в разных базах. Само собою разумеется, что матрица, задающая линейное преобразование, зависит от выбора базы. Покажем, какова связь между матрицами, задающими в разных базах одно и то же линейное преобразование.

Пусть даны базы e и e' с матрицей перехода T,

$$e'=Te, \tag{11}$$

и пусть линейное преобразование φ задается в этих базах соответственно матрицами A и A',

$$e\varphi=Ae, \quad e'\varphi=A'e'. \tag{12}$$

Второе из равенств (12) приводит, ввиду (11), к равенству

$$(Te)\,\varphi=A'\,(Te).$$

Однако

$$(Te)\,\varphi=T\,(e\varphi).$$

Действительно, если $(\tau_{i1},\ \tau_{i2},\ \ldots,\ \tau_{in})$ — i-я строка матрицы T, то

$$(\tau_{i1}e_1+\tau_{i2}e_2+\ldots+\tau_{in}e_n)\,\varphi=\tau_{i1}\,(e_1\varphi)+\tau_{i2}\,(e_2\varphi)+\ldots+\tau_{in}\,(e_n\varphi).$$

Таким образом, ввиду (12),

$$(Te)\,\varphi=T\,(e\varphi)=T\,(Ae)=(TA)\,e,$$
$$A'\,(Te)=(A'T)\,e,$$

т. е.

$$(TA)\,e=(A'T)\,e.$$

Если хотя бы для одного i, $1\leqslant i\leqslant n$, i-я строка матрицы TA будет отлична от i-й строки матрицы $A'T$, то две различные линейные комбинации векторов e_1, e_2, $\ldots$, e_n окажутся равными друг другу, что противоречит линейной независимости базы e. Таким образом,

$$TA=A'T,$$

откуда, ввиду невырожденности матрицы перехода T,

$$A' = TAT^{-1}, \quad A = T^{-1}A'T. \tag{13}$$

Заметим, что квадратные матрицы B и C называются *подобными*, если они связаны равенством

$$C = Q^{-1}BQ,$$

где Q — некоторая невырожденная матрица. При этом говорят, что матрица C получена из матрицы B *трансформированием* матрицей Q.

Доказанные выше равенства (13) можно сформулировать, таким образом, в виде следующей важной теоремы:

Матрицы, задающие одно и то же линейное преобразование в разных базах, подобны между собой. При этом матрица линейного преобразования φ в базе e' получается трансформированием матрицы этого преобразования в базе e матрицей перехода от базы e' к базе e.

Подчеркнем, что если матрица A задает линейное преобразование φ в базе e, то любая матрица B, подобная матрице A,

$$B = Q^{-1}AQ,$$

также задает преобразование φ в некоторой базе, а именно в базе, получающейся из базы e при помощи матрицы перехода Q^{-1}.

Операции над линейными преобразованиями. Сопоставляя каждому линейному преобразованию пространства V_n его матрицу в фиксированной базе, мы получаем, как доказано, взаимно однозначное соответствие между всеми линейными преобразованиями и всеми квадратными матрицами порядка n. Естественно ожидать, что операциям сложения и умножения матриц, а также умножения матрицы на число, будут соответствовать аналогичные операции над линейными преобразованиями.

Пусть в пространстве V_n даны линейные преобразования φ и ψ. Назовем *суммой* этих преобразований преобразование $\varphi + \psi$, определяемое равенством

$$a(\varphi + \psi) = a\varphi + a\psi; \tag{14}$$

оно переводит, следовательно, любой вектор a в сумму его образов при преобразованиях φ и ψ.

Преобразование $\varphi + \psi$ является линейным. Действительно, для любых векторов a и b и любого числа α

$$(a + b)(\varphi + \psi) = (a + b)\varphi + (a + b)\psi =$$
$$= a\varphi + b\varphi + a\psi + b\psi = a(\varphi + \psi) + b(\varphi + \psi);$$

$$(\alpha a)(\varphi + \psi) = (\alpha a)\varphi + (\alpha a)\psi = \alpha(a\varphi) + \alpha(a\psi) =$$
$$= \alpha(a\varphi + a\psi) = \alpha[a(\varphi + \psi)].$$

С другой стороны, назовем *произведением* линейных преобразований φ и ψ преобразование $\varphi\psi$, определяемое равенством

$$a(\varphi\psi) = (a\varphi)\psi, \tag{15}$$

т. е. получающееся в результате последовательного выполнения преобразований φ и ψ.

Преобразование $\varphi\psi$ является линейным:

$$(a+b)(\varphi\psi) = [(a+b)\varphi]\psi = (a\varphi + b\varphi)\psi = \\ = (a\varphi)\psi + (b\varphi)\psi = a(\varphi\psi) + b(\varphi\psi);$$

$$(\alpha a)(\varphi\psi) = [(\alpha a)\varphi]\psi = [\alpha(a\varphi)]\psi = \alpha[(a\varphi)\psi] = \alpha[a(\varphi\psi)].$$

Назовем, наконец, *произведением линейного преобразования φ на число $\varkappa$* преобразование $\varkappa\varphi$, определяемое равенством

$$a(\varkappa\varphi) = \varkappa(a\varphi); \tag{16}$$

образы при преобразовании φ всех векторов умножаются, следовательно, на число $\varkappa$.

Преобразование $\varkappa\varphi$ является линейным:

$$(a+b)(\varkappa\varphi) = \varkappa[(a+b)\varphi] = \varkappa(a\varphi + b\varphi) = \\ = \varkappa(a\varphi) + \varkappa(b\varphi) = a(\varkappa\varphi) + b(\varkappa\varphi);$$

$$(\alpha a)(\varkappa\varphi) = \varkappa[(\alpha a)\varphi] = \varkappa[\alpha(a\varphi)] = \alpha[\varkappa(a\varphi)] = \alpha[a(\varkappa\varphi)].$$

Пусть в базе $e_1, e_2, \ldots, e_n$ преобразования φ и ψ задаются соответственно матрицами $A = (\alpha_{ij})$ и $B = (\beta_{ij})$,

$$e\varphi = Ae, \quad e\psi = Be.$$

Тогда, ввиду (14),

$$e_i(\varphi + \psi) = e_i\varphi + e_i\psi = \sum_{j=1}^{n} \alpha_{ij} e_j + \sum_{j=1}^{n} \beta_{ij} e_j = \sum_{j=1}^{n} (\alpha_{ij} + \beta_{ij}) e_j,$$

т. е.

$$e(\varphi + \psi) = (A + B)e.$$

Таким образом, *матрица суммы линейных преобразований в любой базе равна сумме матриц этих преобразований в той же базе.*

С другой стороны, ввиду (15),

$$e_i(\varphi\psi) = (e_i\varphi)\psi = \left(\sum_{j=1}^{n} \alpha_{ij} e_j\right)\psi = \sum_{j=1}^{n} \alpha_{ij}(e_j\psi) = \\ = \sum_{j=1}^{n} \alpha_{ij}\left(\sum_{k=1}^{n} \beta_{jk} e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} \alpha_{ij}\beta_{jk}\right) e_k,$$

т. е.

$$e(\varphi\psi) = (AB)e.$$

Иными словами, *матрица произведения линейных преобразований в любой базе равна произведению матриц этих преобразований в той же базе.*

Наконец, ввиду (16),

$$e_i(\varkappa\varphi) = \varkappa\,(e_i\varphi) = \varkappa \sum_{j=1}^{n} \alpha_{ij}e_j = \sum_{j=1}^{n} (\varkappa\alpha_{ij})\,e_j,$$

т. е.

$$e(\varkappa\varphi) = (\varkappa A)\,e.$$

Следовательно, *матрица, задающая в некоторой базе произведение линейного преобразования* φ *на число* $\varkappa$, *равна произведению матрицы самого преобразования* φ *в этой базе на число* $\varkappa$.

Из полученных результатов следует, что операции над линейными преобразованиями обладают теми же свойствами, что и операции над матрицами. Так, сложение линейных преобразований коммутативно и ассоциативно, а умножение ассоциативно, но при $n > 1$ не коммутативно. Для линейных преобразований существует однозначное вычитание. Отметим также, что *тождественное преобразование* ε *играет среди линейных преобразований роль единицы, а нулевое преобразование* ω — *роль нуля.* Действительно, в любой базе преобразование ε задается единичной матрицей, а преобразование ω — нулевой матрицей.

§ 32*. Линейные подпространства

Подмножество L линейного пространства V называется *линейным подпространством* этого пространства, если оно само является линейным пространством по отношению к определенным в V операциям сложения векторов и умножения вектора на число. Так, в трехмерном евклидовом пространстве совокупность векторов, выходящих из начала координат и лежащих на некоторой плоскости (или некоторой прямой), проходящей через начало, будет линейным подпространством.

Для того чтобы непустое подмножество L *пространства* V *было его линейным подпространством, достаточно выполнения следующих требований:*

1. *Если векторы* a *и* b *принадлежат к* L, *то в* L *содержится и вектор* $a+b$.

2. *Если вектор* a *принадлежит к* L, *то в* L *содержится и вектор* αa *при любом значении числа* α.

Действительно, ввиду условия 2, множество L содержит нулевой вектор: если вектор a принадлежит к L, то L содержит и вектор $0 \cdot a = 0$. Далее, L вместе со всяким своим вектором a содержит, снова ввиду свойства 2, и противоположный ему вектор $-a = (-1) \cdot a$, а поэтому ввиду свойства 1 к L принадлежит и разность любых двух векторов из L. Что же касается всех остальных

требований, входящих в определение линейного пространства, то они, выполняясь в V, будут выполняться и в L.

Примерами линейных подпространств пространства V могут служить само пространство V, а также множество, состоящее из одного нулевого вектора, — так называемое *нулевое подпространство*. Более интересен следующий пример: берем в пространстве V любую конечную систему векторов

$$a_1,\ a_2,\ \ldots,\ a_r \tag{1}$$

и обозначаем через L множество всех тех векторов, которые являются линейными комбинациями векторов (1). Докажем, что L будет линейным подпространством. В самом деле, если

$$b=\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_r a_r,\quad c=\beta_1 a_1+\beta_2 a_2+\ldots+\beta_r a_r,$$

то

$$b+c=(\alpha_1+\beta_1)\,a_1+(\alpha_2+\beta_2)\,a_2+\ldots+(\alpha_r+\beta_r)\,a_r,$$

т. е. вектор $b+c$ принадлежит к L; к L принадлежит и вектор

$$\gamma b=(\gamma\alpha_1)\,a_1+(\gamma\alpha_2)\,a_2+\ldots+(\gamma\alpha_r)\,a_r$$

при любом числе γ.

Говорят, что это линейное подпространство L *порождено* системой векторов (1); к L принадлежат, в частности, сами векторы (1).

Впрочем, *всякое линейное подпространство конечномерного линейного пространства порождается конечной системой векторов*, так как если оно не является нулевым, то обладает даже конечной базой. Размерность линейного подпространства L не больше размерности n самого пространства V_n, причем равна n лишь при $L=V_n$. Размерностью нулевого подпространства следует считать, конечно, число 0.

Для всякого k, $0<k<n$, в пространстве V_n существуют линейные подпространства размерности k — достаточно взять подпространство, порожденное любой системой из k линейно независимых векторов.

Пусть в пространстве V даны линейные подпространства L_1 и L_2. Совокупность L_0 векторов, принадлежащих как к L_1, так и к L_2, будет, как легко проверить, линейным подпространством; оно называется *пересечением* подпространств L_1 и L_2. С другой стороны, линейным подпространством будет и *сумма* $\bar{L}$ подпространств L_1 и L_2, т. е. совокупность всех тех векторов из V, которые представимы в виде суммы двух слагаемых, одного из L_1, другого из L_2. Если размерности подпространств L_1, L_2, L_0 и $\bar{L}$ суть, соответственно, d_1, d_2, d_0 и $\bar{d}$, то имеет место следующая формула:

$$\bar{d}=d_1+d_2-d_0, \tag{2}$$

т. е. *размерность суммы двух подпространств равна сумме размерностей этих подпространств минус размерность их пересечения.*

Для доказательства берем произвольную базу

$$a_1, a_2, \ldots, a_{d_0} \tag{3}$$

подпространства L_0 и дополняем ее до базы

$$a_1, a_2, \ldots, a_{d_0}, b_{d_0+1}, \ldots, b_{d_1} \tag{4}$$

подпространства L_1 и до базы

$$a_1, a_2, \ldots, a_{d_0}, c_{d_0+1}, \ldots, c_{d_2} \tag{5}$$

подпространства L_2. Легко видеть, используя определение подпространства $\bar{L}$, что это подпространство порождается системой векторов

$$a_1, a_2, \ldots, a_{d_0}, b_{d_0+1}, \ldots, b_{d_1}, c_{d_0+1}, \ldots, c_{d_2}. \tag{6}$$

Формула (2) будет, следовательно, доказана, если мы докажем линейную независимость системы (6).

Пусть имеет место равенство

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{d_0} a_{d_0} + \beta_{d_0+1} b_{d_0+1} + \ldots + \beta_{d_1} b_{d_1} + \\ + \gamma_{d_0+1} c_{d_0+1} + \ldots + \gamma_{d_2} c_{d_2} = 0$$

с некоторыми числовыми коэффициентами. Тогда

$$d = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_{d_0} a_{d_0} + \beta_{d_0+1} b_{d_0+1} + \ldots + \beta_{d_1} b_{d_1} = \\ = -\gamma_{d_0+1} c_{d_0+1} - \ldots - \gamma_{d_2} c_{d_2}. \tag{7}$$

Левая часть этого равенства содержится в L_1, правая — в L_2, поэтому вектор d, равный как левой, так и правой части этого равенства, принадлежит к L_0 и, следовательно, линейно выражается через базу (3). Правая часть равенства (7) показывает, однако, что вектор d линейно выражается и через векторы $c_{d_0+1}, \ldots, c_{d_2}$. Отсюда, ввиду линейной независимости системы (5), вытекает, что все коэффициенты $\gamma_{d_0+1}, \ldots, \gamma_{d_2}$ равны нулю, т. е. $d=0$, а тогда, ввиду линейной независимости системы (4), все коэффициенты $\alpha_1, \ldots, \alpha_{d_0}, \beta_{d_0+1}, \ldots, \beta_{d_1}$ также равны нулю. Этим доказана линейная независимость системы (6).

Читателю предлагается проверить, что наше доказательство сохраняет силу и в том случае, когда подпространство L_0 является нулевым, т. е. $d_0=0$.

Область значений и ядро линейного преобразования. Пусть в линейном пространстве V_n задано линейное преобразование φ. Если L — любое линейное подпространство пространства V_n, то *совокупность $L\varphi$ образов всех векторов из L при преобразовании φ также будет линейным подпространством*, как немедленно вытекает из определений линейного

подпространства и линейного преобразования. В частности, линейным подпространством будет и совокупность $V_n\varphi$ образов всех векторов пространства V_n; она называется *областью значений* преобразования φ.

Найдем размерность области значений. Для этого заметим, что так как все матрицы, задающие преобразование φ в разных базах, подобны между собой, то, ввиду последней теоремы из § 14, все они имеют один и тот же ранг. Это число можно назвать, следовательно, *рангом* линейного преобразования φ.

Размерность области значений линейного преобразования φ равна рангу этого преобразования.

В самом деле, пусть φ задается в базе $e_1, e_2, \ldots, e_n$ матрицей A. Подпространство $V_n\varphi$ порождается векторами

$$e_1\varphi,\ e_2\varphi,\ \ldots,\ e_n\varphi \tag{8}$$

и поэтому базой подпространства $V_n\varphi$ будет служить, в частности, любая максимальная линейно независимая подсистема системы (8). Однако максимальное число линейно независимых векторов в системе (8) равно максимальному числу линейно независимых строк матрицы A, т. е. равно рангу этой матрицы. Теорема доказана.

Мы знаем, что при линейном преобразовании φ нулевой вектор переходит в самого себя. Совокупность $N(\varphi)$ всех векторов пространства V_n, отображающихся при φ в нулевой вектор, будет, следовательно, непустой и является, очевидно, линейным подпространством. Это подпространство называется *ядром* преобразования φ, а его размерность — *дефектом* этого преобразования.

Для любого линейного преобразования φ пространства V_n сумма ранга и дефекта этого преобразования равна размерности n всего пространства.

Действительно, если r — ранг преобразования φ, то подпространство $V_n\varphi$ обладает базой из r векторов

$$a_1,\ a_2,\ \ldots,\ a_r. \tag{9}$$

В пространстве V_n можно выбрать такие векторы

$$b_1,\ b_2,\ \ldots,\ b_r, \tag{10}$$

что

$$b_i\varphi = a_i,\quad i = 1,\ 2,\ \ldots,\ r;$$

выбор векторов (10) не является, понятно, однозначным. Если бы некоторая нетривиальная линейная комбинация векторов (10) отображалась преобразованием φ в нуль, в частности, если бы векторы (10) были линейно зависимыми, то векторы (9) оказались бы сами линейно зависимыми против предположения. Поэтому линейное подпространство L, порожденное векторами (10), имеет размерность r, а его пересечение с подпространством $N(\varphi)$ равно нулю.

С другой стороны, сумма подпространств L и $N(\varphi)$ совпадает со всем пространством V_n. Действительно, если c — любой вектор пространства, то вектор $d = c\varphi$ принадлежит, конечно, к подпространству $V_n\varphi$. Тогда в подпространстве L найдется такой вектор b, что

$$b\varphi = d$$

— вектор b записывается через систему (10) с теми же коэффициентами, с какими вектор d записывается через базу (9). Отсюда

$$c = b + (c - b),$$

причем вектор $c - b$ содержится в подпространстве $N(\varphi)$, так как

$$(c - b)\,\varphi = c\varphi - b\varphi = d - d = 0.$$

Из полученных результатов и доказанной выше формулы (2) вытекает утверждение теоремы.

Невырожденные линейные преобразования. Линейное преобразование φ линейного пространства V_n называется *невырожденным*, если оно удовлетворяет любому из следующих условий, равносильность которых немедленно вытекает из доказанных выше теорем:

1. Ранг преобразования φ равен n.

2. Областью значений преобразования φ служит все пространство V_n.

3. Дефект преобразования φ равен нулю.

Для невырожденных линейных преобразований можно указать также много других определений, равносильных указанным выше, в частности определения 4—6.

4. Различные векторы пространства V_n имеют при преобразовании φ различные образы.

Действительно, если преобразование φ обладает свойством 4, то ядро этого преобразования состоит лишь из нулевого вектора, т. е. выполняется и свойство 3. Если же векторы a и b таковы, что $a \neq b$, но $a\varphi = b\varphi$, то $a - b \neq 0$, но $(a - b)\varphi = 0$, т. е. свойство 3 не выполняется.

Из 2 и 4 вытекает

5. Преобразование φ является взаимно однозначным отображением пространства V_n на все это пространство.

Из 5 следует, что для невырожденного линейного преобразования φ существует *обратное преобразование* φ^{-1}, переводящее всякий вектор $a\varphi$ в вектор a,

$$(a\varphi)\,\varphi^{-1} = a.$$

Преобразование φ^{-1} *будет линейным*, так как

$$(a\varphi + b\varphi)\,\varphi^{-1} = [(a + b)\,\varphi]\,\varphi^{-1} = a + b,$$

$$[\alpha\,(a\varphi)]\varphi^{-1} = [(\alpha a)\varphi]\,\varphi^{-1} = \alpha a.$$

Из определения преобразования φ^{-1} вытекает, что

$$\varphi\varphi^{-1}=\varphi^{-1}\varphi=\varepsilon; \tag{11}$$

равенства (11) могут сами рассматриваться как определение обратного преобразования. Отсюда и из последних результатов предшествующего параграфа следует, что *если невырожденное линейное преобразование φ задается в некоторой базе матрицей A, невырожденной ввиду свойства 1, то преобразование φ^{-1} задается в этой базе матрицей A^{-1}.*

Мы приходим, таким образом, к следующему определению невырожденного линейного преобразования:

6. Для преобразования φ существует обратное линейное преобразование φ^{-1}.

§ 33. Характеристические корни и собственные значения

Пусть $A=(\alpha_{ij})$ — квадратная матрица порядка n с действительными элементами. Пусть, с другой стороны, λ — некоторое неизвестное. Тогда матрица $A-\lambda E$, где E — единичная матрица порядка n, называется *характеристической матрицей* матрицы A. Так как в матрице λE по главной диагонали стоит λ, все же остальные элементы равны нулю, то

$$A-\lambda E=\begin{pmatrix} \alpha_{11}-\lambda & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn}-\lambda \end{pmatrix}.$$

Определитель матрицы $A-\lambda E$ будет многочленом от λ, притом степени n. В самом деле, произведение элементов, стоящих на главной диагонали, будет многочленом от λ со старшим членом $(-1)^n\lambda^n$, все же остальные члены определителя не содержат по меньшей мере двух из числа элементов, стоящих на главной диагонали, и поэтому их степень относительно λ не превосходит $n-2$. Коэффициенты этого многочлена можно было бы легко найти. Так, коэффициент при λ^{n-1} равен $(-1)^{n-1}(\alpha_{11}+\alpha_{22}+\dots+\alpha_{nn})$, а свободный член совпадает с определителем матрицы A.

Многочлен n-й степени $|A-\lambda E|$ называется *характеристическим многочленом* матрицы A, а его корни, которые могут быть как действительными, так и комплексными, называются *характеристическими корнями* этой матрицы.

Подобные матрицы обладают одинаковыми характеристическими многочленами и, следовательно, одинаковыми характеристическими корнями.

Пусть, в самом деле,

$$B = Q^{-1}AQ.$$

Тогда, учитывая, что матрица λE перестановочна с матрицей Q, а $|Q^{-1}| = |Q|^{-1}$, получаем:

$$|B - \lambda E| = |Q^{-1}AQ - \lambda E| = |Q^{-1}(A - \lambda E)Q| = \\ = |Q|^{-1} \cdot |A - \lambda E| \cdot |Q| = |A - \lambda E|,$$

что и требовалось доказать.

Из этого результата вытекает, ввиду доказанной в § 31 теоремы о связи между матрицами, задающими линейное преобразование в разных базах, что *хотя линейное преобразование* φ *может задаваться в разных базах различными матрицами, однако все эти матрицы имеют один и тот же набор характеристических корней*. Эти корни можно называть поэтому *характеристическими корнями самого преобразования* φ. Весь набор этих характеристических корней, причем каждый корень берется с той кратностью, какую он имеет в характеристическом многочлене, называется *спектром* линейного преобразования φ.

Характеристические корни играют при изучении линейных преобразований очень большую роль. Читатель много раз будет иметь возможность в этом убедиться. Одно из применений характеристических корней мы сейчас укажем.

Пусть в действительном линейном пространстве V_n задано линейное преобразование φ. Если вектор b, отличный от нуля, переводится преобразованием φ в вектор, пропорциональный самому b,

$$b\varphi = \lambda_0 b, \qquad (1)$$

где λ_0 — некоторое действительное число, то вектор b называется *собственным вектором* преобразования φ, а число λ_0 — *собственным значением* этого преобразования, причем говорят, что собственный вектор b *относится* к собственному значению λ_0.

Заметим, что так как $b \neq 0$, то число λ_0, удовлетворяющее условию (1), определяется для вектора b однозначно. Подчеркнем, далее, что нулевой вектор не считается собственным вектором преобразования φ, хотя он удовлетворяет условию (1), притом для любого λ_0.

Вращение евклидовой плоскости вокруг начала координат на угол, не являющийся кратным π, служит примером линейного преобразования, не имеющего собственных векторов. Примером другого крайнего случая является растяжение плоскости, при котором все векторы, выходящие из начала координат, растягиваются, скажем, в пять раз. Это будет линейное преобразование, причем все ненулевые векторы плоскости будут для него собственными; все они относятся к собственному значению 5.

Действительные характеристические корни линейного преобразования φ, *если они существуют, и только они служат собственными значениями этого преобразования.*

Пусть, в самом деле, преобразование φ имеет в базе $e_1, e_2, \ldots, e_n$ матрицу $A=(\alpha_{ij})$ и пусть вектор

$$b=\sum_{i=1}^{n}\beta_i e_i$$

является собственным вектором преобразования φ,

$$b\varphi=\lambda_0 b. \tag{2}$$

Как доказано в § 31,

$$b\varphi=[(\beta_1, \beta_2, \ldots, \beta_n)A]e. \tag{3}$$

Равенства (2) и (3) приводят к системе равенств

$$\begin{gathered}\beta_1\alpha_{11}+\beta_2\alpha_{21}+\ldots+\beta_n\alpha_{n1}=\lambda_0\beta_1,\\ \beta_1\alpha_{12}+\beta_2\alpha_{22}+\ldots+\beta_n\alpha_{n2}=\lambda_0\beta_2,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \beta_1\alpha_{1n}+\beta_2\alpha_{2n}+\ldots+\beta_n\alpha_{nn}=\lambda_0\beta_n.\end{gathered}\tag{4}$$

Так как $b\neq 0$, то не все числа $\beta_1, \beta_2, \ldots, \beta_n$ равны нулю. Таким образом, ввиду (4), система линейных однородных уравнений

$$\begin{gathered}(\alpha_{11}-\lambda_0)x_1+\alpha_{21}x_2+\ldots+\alpha_{n1}x_n=0,\\ \alpha_{12}x_1+(\alpha_{22}-\lambda_0)x_2+\ldots+\alpha_{n2}x_n=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \alpha_{1n}x_1+\alpha_{2n}x_2+\ldots+(\alpha_{nn}-\lambda_0)x_n=0\end{gathered}\tag{5}$$

обладает ненулевым решением, а поэтому ее определитель равен нулю,

$$\begin{vmatrix}\alpha_{11}-\lambda_0, & \alpha_{21}, & \ldots, & \alpha_{n1}\\ \alpha_{12}, & \alpha_{22}-\lambda_0, & \ldots, & \alpha_{n2}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{1n}, & \alpha_{2n}, & \ldots, & \alpha_{nn}-\lambda_0\end{vmatrix}=0. \tag{6}$$

Транспонируя, получаем

$$|A-\lambda_0 E|=0, \tag{7}$$

т. е. собственное значение λ_0 на самом деле оказалось характеристическим корнем матрицы A и, следовательно, линейного преобразования φ, притом, понятно, действительным.

Обратно, пусть λ_0 будет любым действительным характеристическим корнем преобразования φ и, следовательно, матрицы A. Тогда имеет место равенство (7), а поэтому и получающееся из него транспонированием равенство (6). Отсюда следует, что система линейных однородных уравнений (5) обладает ненулевым решением,

притом даже действительным, так как все коэффициенты этой системы действительны. Если это решение обозначим через

$$(\beta_1, \beta_2, \ldots, \beta_n), \tag{8}$$

то имеют место равенства (4). Обозначим через b вектор пространства V_n, имеющий в базе $e_1, e_2, \ldots, e_n$ строку координат (8); ясно, что $b \neq 0$. Тогда справедливо равенство (3), а из (4) и (3) следует (2). Вектор b оказался, таким образом, собственным вектором преобразования φ, относящимся к собственному значению λ_0. Теорема доказана.

Заметим, что если бы мы рассматривали комплексное линейное пространство, то требование действительности характеристического корня было бы излишним, т. е. была бы доказана теорема: *характеристические корни линейного преобразования комплексного линейного пространства и только они служат собственными значениями этого преобразования*. Отсюда следует, что *в комплексном линейном пространстве всякое линейное преобразование обладает собственными векторами.*

Возвращаясь к рассматриваемому нами действительному случаю, отметим, что совокупность собственных векторов линейного преобразования φ, относящихся к собственному значению λ_0, совпадает с совокупностью ненулевых действительных решений системы линейных однородных уравнений (5). Отсюда следует, что *совокупность собственных векторов линейного преобразования φ, относящихся к собственному значению λ_0, будет, после добавления к ней нулевого вектора, линейным подпространством пространства V_n*. В самом деле, из доказанного в § 12 вытекает, что *совокупность (действительных) решений любой системы линейных однородных уравнений от n неизвестных будет линейным подпространством пространства V_n.*

Линейные преобразования с простым спектром. Во многих случаях оказывается необходимым знать, может ли данное линейное преобразование φ иметь в некоторой базе диагональную матрицу. На самом деле далеко не всякое линейное преобразование может быть задано диагональной матрицей. Необходимые и достаточные условия для этого будут указаны в § 61, а сейчас мы хотим привести одно достаточное условие.

Докажем сначала следующие вспомогательные результаты:

Линейное преобразование φ тогда и только тогда задается в базе $e_1, e_2, \ldots, e_n$ диагональной матрицей, если все векторы этой базы являются собственными векторами преобразования φ.

Действительно, равенство

$$e_i\varphi = \lambda_i e_i$$

равносильно тому, что в i-й строке матрицы, задающей преобразование φ в указанной базе, равны нулю все элементы, стоящие

вне главной диагонали, а на главной диагонали (т. е. на i-м месте) стоит число λ_i.

Собственные векторы $b_1, b_2, \ldots, b_k$ *линейного преобразования* φ, *относящиеся к различным собственным значениям, составляют линейно независимую систему.*

Будем доказывать это утверждение индукцией по k, так как при $k=1$ оно справедливо — один собственный вектор, будучи отличным от нуля, составляет линейно независимую систему. Пусть

$$b_i\varphi = \lambda_i b_i, \quad i = 1, 2, \ldots, k,$$

и

$$\lambda_i \neq \lambda_j \text{ при } i \neq j.$$

Если существует линейная зависимость

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_k b_k = 0, \tag{9}$$

где, например, $\alpha_1 \neq 0$, то, применяя к обеим частям равенства (9) преобразование φ, получим

$$\alpha_1 \lambda_1 b_1 + \alpha_2 \lambda_2 b_2 + \ldots + \alpha_k \lambda_k b_k = 0.$$

Вычитая отсюда равенство (9), умноженное на λ_k, получаем

$$\alpha_1(\lambda_1 - \lambda_k) b_1 + \alpha_2(\lambda_2 - \lambda_k) b_2 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) b_{k-1} = 0,$$

что дает нетривиальную линейную зависимость между векторами $b_1, b_2, \ldots, b_{k-1}$, так как $\alpha_1(\lambda_1 - \lambda_k) \neq 0$.

Говорят, что линейное преобразование φ действительного линейного пространства V_n имеет *простой спектр*, если все его характеристические корни действительны и различны. Преобразование φ имеет, следовательно, n различных собственных значений, а поэтому, по доказанной теореме, в пространстве V_n существует база, составленная из собственных векторов этого преобразования. Таким образом, *всякое линейное преобразование с простым спектром может быть задано диагональной матрицей.*

Переходя от линейного преобразования к матрицам, его задающим, мы получаем следующий результат:

Всякая матрица, все характеристические корни которой действительны и различны, подобна диагональной матрице или, как говорят, *такая матрица приводится к диагональному виду.*

ГЛАВА ВОСЬМАЯ

ЕВКЛИДОВЫ ПРОСТРАНСТВА

§ 34. Определение евклидова пространства. Ортонормированные базы

Понятие n-мерного линейного пространства далеко не в полной мере обобщает понятия плоскости или трехмерного евклидова пространства — в n-мерном случае при $n>3$ не определены ни длина вектора, ни угол между векторами, и поэтому невозможно развитие той богатой геометрической теории, которая хорошо знакома читателю для $n=2$ и $n=3$. Оказывается, что положение может быть исправлено, притом следующим путем.

Из курса аналитической геометрии известно, что и в плоскости, и в трехмерном пространстве можно ввести понятие скалярного умножения векторов. Оно определяется при помощи длин векторов и угла между ними, но, как оказывается, и длина вектора, и угол между векторами в свою очередь могут быть выражены через скалярные произведения. Мы определим поэтому в любом n-мерном линейном пространстве понятие скалярного умножения, причем определим аксиоматически, при помощи некоторых свойств, которыми, как хорошо известно, скалярное умножение векторов плоскости или трехмерного пространства на самом деле обладает. При этом, учитывая те непосредственные цели, ради которых этот раздел включен в курс высшей алгебры, вводить определения длины вектора и угла между векторами мы не станем. Читателя, интересующегося построением геометрии в n-мерных пространствах, мы отсылаем к специальной литературе, в первую очередь к более полным книгам по линейной алгебре.

Отметим, что всюду в этой главе, кроме конца настоящего параграфа, рассматриваются действительные линейные пространства.

Будем говорить, что в n-мерном действительном линейном пространстве V_n определено *скалярное умножение*, если всякой паре векторов a, b поставлено в соответствие действительное число, обозначаемое символом (a, b) и называемое *скалярным произведением* векторов a и b, причем выполняются следующие условия

(здесь a, b, c — любые векторы пространства V_n, α — любое действительное число):

I. $(a,\ b)=(b,\ a)$.

II. $(a+b,\ c)=(a,\ c)+(b,\ c)$.

III. $(\alpha a,\ b)=\alpha(a,\ b)$.

IV. Если $a\neq 0$, то скалярный квадрат вектора a строго положителен,

$$(a,\ a)>0.$$

Отметим, что из III при $\alpha=0$ следует равенство

$$(0,\ b)=0, \qquad (1)$$

т. е. *скалярное произведение нулевого вектора на любой вектор b равно нулю*; равен нулю, в частности, скалярный квадрат нулевого вектора.

Из II и III немедленно вытекает следующая формула для скалярного произведения линейных комбинаций двух систем векторов:

$$\left(\sum_{i=1}^{k}\alpha_i a_i,\ \sum_{j=1}^{l}\beta_j b_j\right)=\sum_{i=1}^{k}\sum_{j=1}^{l}\alpha_i\beta_j(a_i,\ b_j). \qquad (2)$$

Если в n-мерном линейном пространстве определено скалярное умножение, то это пространство называется n-мерным *евклидовым пространством*.

При любом n в n-мерном линейном пространстве V_n можно определить скалярное умножение, т. е. можно превратить это пространство в евклидово.

В самом деле, возьмем в пространстве V_n любую базу e_1, e_2, ..., e_n. Если

$$a=\sum_{i=1}^{n}\alpha_i e_i, \quad b=\sum_{i=1}^{n}\beta_i e_i,$$

то положим

$$(a,\ b)=\sum_{i=1}^{n}\alpha_i\beta_i. \qquad (3)$$

Легко проверяется, что условия I — IV будут выполнены, т. е. равенство (1) определяет в пространстве V_n скалярное умножение.

Мы видим, что в n-мерном линейном пространстве скалярное умножение можно задать, вообще говоря, многими различными способами — определение (3) зависит, понятно, от выбора базы, а мы пока не знаем, кроме того, нельзя ли ввести скалярное умножение и каким-либо принципиально иным способом. Нашей ближайшей целью является обозрение всех возможных способов пре-

вращения n-мерного линейного пространства в евклидово пространство и установление того, что в некотором смысле для всякого n существует одно-единственное n-мерное евклидово пространство.

Пусть дано произвольное n-мерное евклидово пространство E_n, т. е. в n-мерном линейном пространстве произвольным способом введено скалярное умножение. Векторы a и b называются *ортогональными*, если их скалярное произведение равно нулю,

$$(a, b) = 0.$$

Из (1) следует, что нулевой вектор ортогонален к любому вектору; могут существовать, однако, и ненулевые ортогональные векторы.

Система векторов называется *ортогональной системой*, если все векторы этой системы попарно ортогональны между собой.

Всякая ортогональная система ненулевых векторов линейно независима.

Пусть, в самом деле, в E_n дана система векторов $a_1, a_2, \ldots, a_k$, причем $a_i \neq 0$, $i = 1, 2, \ldots, k$, и

$$(a_i, a_j) = 0 \quad \text{при} \quad i \neq j. \tag{4}$$

Если

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0,$$

то, скалярно умножая обе части этого равенства на вектор a_i $1 \leqslant i \leqslant k$, получаем ввиду (1), (2) и (4):

$$0 = (0, a_i) = (\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k, a_i) =$$
$$= \alpha_1 (a_1, a_i) + \alpha_2 (a_2, a_i) + \ldots + \alpha_k (a_k, a_i) = \alpha_i (a_i, a_i).$$

Отсюда, так как $(a_i, a_i) > 0$ по IV, вытекает $\alpha_i = 0$, $i = 1, 2, \ldots, k$, что и требовалось доказать.

Сейчас будет описан процесс ортогонализации, т. е. некоторый способ перехода от любой линейно независимой системы из k векторов

$$a_1, a_2, \ldots, a_k \tag{5}$$

евклидова пространства E_n к ортогональной системе, также состоящей из k ненулевых векторов; эти векторы будут обозначены через $b_1, b_2, \ldots, b_k$.

Положим $b_1 = a_1$, т. е. первый вектор системы (5) войдет и в строящуюся нами ортогональную систему. Положим, далее,

$$b_2 = \alpha_1 b_1 + a_2.$$

Так как $b_1 = a_1$, а векторы a_1 и a_2 линейно независимы, то вектор b_2 отличен от нуля при любом числе α_1. Подберем это число из условия, что вектор b_2 должен быть ортогонален к вектору b_1:

$$0 = (b_1, b_2) = (b_1, \alpha_1 b_1 + a_2) = \alpha_1 (b_1, b_1) + (b_1, a_2),$$

откуда, ввиду IV,

$$\alpha_1 = -\frac{(b_1, a_2)}{(b_1, b_1)}.$$

Пусть уже построена ортогональная система ненулевых векторов $b_1, b_2, \ldots, b_l$; дополнительно предположим, что для всякого i, $1 \leqslant i \leqslant l$, вектор b_i является линейной комбинацией векторов a_1, $a_2, \ldots, a_i$. Это предположение будет выполняться тогда и для вектора b_{l+1}, если он будет выбран в виде

$$b_{l+1} = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_l b_l + a_{l+1}.$$

Вектор b_{l+1} будет при этом отличен от нуля, так как система (5) линейно независимая, а вектор a_{l+1} не входит в записи векторов $b_1, b_2, \ldots, b_l$. Коэффициенты α_i, $i = 1, 2, \ldots, l$, подберем из условия, что вектор b_{l+1} должен быть ортогонален ко всем векторам b_i, $i = 1, 2, \ldots, l$:

$$0 = (b_i, b_{l+1}) = (b_i, \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_l b_l + a_{l+1}) =$$
$$= \alpha_1 (b_i, b_1) + \alpha_2 (b_i, b_2) + \ldots + \alpha_l (b_i, b_l) + (b_i, a_{l+1});$$

отсюда, так как векторы $b_1, b_2, \ldots, b_l$ ортогональны между собой,

$$\alpha_i (b_i, b_i) + (b_i, a_{l+1}) = 0,$$

т. е.

$$\alpha_i = -\frac{(b_i, a_{l+1})}{(b_i, b_i)}, \qquad i = 1, 2, \ldots, l.$$

Продолжая этот процесс, мы построим искомую ортогональную систему $b_1, b_2, \ldots, b_k$.

Применяя процесс ортогонализации к произвольной базе пространства E_n, мы получим ортогональную систему из n ненулевых векторов, т. е., так как эта система по доказанному линейно независима, *ортогональную базу*. При этом, используя замечание, сделанное в связи с первым шагом процесса ортогонализации, а также учитывая, что всякий ненулевой вектор можно включить в некоторую базу пространства, можно сформулировать даже следующее утверждение:

Всякое евклидово пространство обладает ортогональными базами, причем любой ненулевой вектор этого пространства входит в состав некоторой ортогональной базы.

В дальнейшем важную роль будет играть один специальный вид ортогональных баз; базы этого вида соответствуют прямоугольным декартовым системам координат, используемым в аналитической геометрии.

Назовем вектор b *нормированным*, если его скалярный квадрат равен единице,

$$(b, b) = 1.$$

Если $a \neq 0$, откуда $(a, a) > 0$, то *нормированием* вектора a называется переход к вектору

$$b = \frac{1}{\sqrt{(a, a)}} a.$$

Вектор b будет нормированным, так как

$$(b, b) = \left(\frac{1}{\sqrt{(a, a)}} a, \frac{1}{\sqrt{(a, a)}} a\right) = \left(\frac{1}{\sqrt{(a, a)}}\right)^2 (a, a) = 1.$$

База $e_1, e_2, \ldots, e_n$ евклидова пространства E_n называется *ортонормированной*, если она ортогональна, а все ее векторы нормированы, т. е.

$$\begin{aligned} &(e_i, e_j) = 0 \text{ при } i \neq j, \\ &(e_i, e_i) = 1, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{6}$$

Всякое евклидово пространство обладает ортонормированными базами.

Для доказательства достаточно взять любую ортогональную базу и нормировать все ее векторы. База останется при этом ортогональной, так как при любых α и β из $(a, b) = 0$ следует

$$(\alpha a, \beta b) = \alpha\beta(a, b) = 0.$$

База $e_1, e_2, \ldots, e_n$ евклидова пространства E_n тогда и только тогда будет ортонормированной, если скалярное произведение любых двух векторов пространства равно сумме произведений соответственных координат этих векторов в указанной базе, т. е. из

$$a = \sum_{i=1}^{n} \alpha_i e_i, \quad b = \sum_{j=1}^{n} \beta_j e_j \tag{7}$$

следует

$$(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i. \tag{8}$$

Действительно, если для нашей базы выполняются равенства (6), то

$$(a, b) = \left(\sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j\right) = \sum_{i, j=1}^{n} \alpha_i \beta_j (e_i, e_j) = \sum_{i=1}^{n} \alpha_i \beta_i.$$

Обратно, если наша база такова, что для любых векторов a и b, записанных в этой базе в виде (7), справедливо равенство (8), то, беря в качестве a и b любые два вектора этой базы e_i и e_j, различные или одинаковые, мы из (8) выведем равенства (6).

Сопоставляя полученный сейчас результат с изложенным ранее доказательством существования n-мерных евклидовых пространств для любого n, можно высказать следующее утверждение: *если в n-мерном линейном пространстве V_n выбрана произвольная база,*

то в V_n можно так задать скалярное умножение, что в полученном евклидовом пространстве выбранная база будет одной из ортонормированных баз.

Изоморфизм евклидовых пространств. Евклидовы пространства E и E' называются *изоморфными*, если между векторами этих пространств можно установить такое взаимно однозначное соответствие, что выполняются следующие требования:

1) это соответствие является изоморфным соответствием между E и E', рассматриваемыми как линейные пространства (см. § 29);

2) при этом соответствии сохраняется скалярное произведение; иными словами, если образами векторов a и b из E служат соответственно векторы a' и b' из E', то

$$(a, b) = (a', b'). \tag{9}$$

Из условия 1) сразу следует, что *изоморфные евклидовы пространства имеют одну и ту же размерность.* Докажем обратное утверждение:

Любые евклидовы пространства E и E', имеющие одну и ту же размерность n, изоморфны между собой.

В самом деле, выберем в пространствах E и E' ортонормированные базы

$$e_1, e_2, \ldots, e_n \tag{10}$$

и, соответственно,

$$e'_1, e'_2, \ldots, e'_n. \tag{11}$$

Ставя в соответствие всякому вектору

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

из E вектор

$$a' = \sum_{i=1}^{n} \alpha_i e'_i$$

из E', имеющий в базе (11) те же координаты, что и вектор a в базе (10), мы получим, очевидно, изоморфное соответствие между линейными пространствами E и E'. Покажем, что выполняется и равенство (9): если

$$b = \sum_{i=1}^{n} \beta_i e_i, \quad b' = \sum_{i=1}^{n} \beta_i e'_i,$$

то, в силу (8) — учесть ортонормированность баз (10) и (11)! —

$$(a, b) = \sum_{i=1}^{n} \alpha_i \beta_i = (a', b').$$

Естественно изоморфные евклидовы пространства не считать различными. Поэтому для всякого n существует единственное n-мерное

евклидово пространство в том же смысле, в каком для всякого n существует единственное n-мерное действительное линейное пространство.

На случай **комплексных** линейных пространств понятия и результаты настоящего параграфа переносятся следующим образом. Комплексное линейное пространство называется *унитарным пространством*, если в нем задано скалярное умножение, причем (a, b) будет, вообще говоря, комплексным числом; при этом должны выполняться аксиомы II—IV (в формулировке последней аксиомы следует подчеркнуть, что скалярный квадрат ненулевого вектора действителен и строго положителен), а аксиома I заменяется аксиомой

I' $$(a, b)=\overline{(b, a)},$$

где черта обозначает, как обычно, переход к сопряженному комплексному числу.

Скалярное умножение уже не будет, следовательно, коммутативным. Тем не менее, равенство, симметричное аксиоме II, остается справедливым,

II' $$(a, b+c)=(a, b)+(a, c),$$

так как

$$(a, b+c)=\overline{(b+c, a)}=\overline{(b, a)+(c, a)}=\overline{(b, a)}+\overline{(c, a)}=(a, b)+(a, c).$$

С другой стороны,

III' $$(a, \alpha b)=\overline{\alpha}\,(a, b),$$

так как

$$(a, \alpha b)=\overline{(\alpha b, a)}=\overline{\alpha (b, a)}=\overline{\alpha}\,\overline{(b, a)}=\overline{\alpha}\,(a, b).$$

Понятия ортогональности и ортонормированной системы векторов переносятся на случай унитарных пространств без всяких изменений. Как и выше, доказывается существование ортонормированных баз во всяком конечномерном унитарном пространстве. При этом, однако, если $e_1, e_2, \ldots, e_n$ — ортонормированная база и векторы a, b имеют в этой базе записи (7), то

$$(a, b)=\sum_{i=1}^{n} \alpha_i \overline{\beta_i}.$$

Результаты дальнейших параграфов настоящей главы также можно было бы перенести с евклидовых на унитарные пространства. Мы не будем этого делать и отошлем интересующегося читателя к специальным книгам по линейной алгебре.

§ 35. Ортогональные матрицы, ортогональные преобразования

Пусть дано действительное линейное преобразование n неизвестных:

$$x_i=\sum_{k=1}^{n} q_{ik} y_k, \quad i=1, 2, \ldots, n; \tag{1}$$

матрицу этого преобразования обозначим через Q. Это преобразование переводит сумму квадратов неизвестных $x_1, x_2, \ldots, x_n$, т. е. квадратичную форму $x_1^2+x_2^2+\ldots+x_n^2$, являющуюся нормальным

видом положительно определенных квадратичных форм (см. § 28), в некоторую квадратичную форму от неизвестных $y_1, y_2, \ldots, y_n$. Случайно эта новая квадратичная форма сама может оказаться суммой квадратов неизвестных $y_1, y_2, \ldots, y_n$, т. е. может иметь место равенство

$$x_1^2+x_2^2+\ldots+x_n^2=y_1^2+y_2^2+\ldots+y_n^2, \qquad (2)$$

тождественное после замены неизвестных $x_1, x_2, \ldots, x_n$ их выражениями (1). Линейное преобразование неизвестных (1), обладающее этим свойством, т. е., как говорят, оставляющее сумму квадратов неизвестных инвариантной, называется *ортогональным преобразованием неизвестных,* а его матрица Q — *ортогональной матрицей.*

Существует много других определений ортогонального преобразования и ортогональной матрицы, эквивалентных приведенным выше. Укажем некоторые из них, необходимые для дальнейшего.

Мы знаем из § 26 закон, по которому преобразуется матрица квадратичной формы при выполнении линейного преобразования неизвестных. Применяя его к нашему случаю и учитывая, что матрицей квадратичной формы, являющейся суммой квадратов всех неизвестных, служит единичная матрица E, мы получим, что равенство (2) равносильно матричному равенству

$$Q'EQ=E,$$

т. е.

$$Q'Q=E. \qquad (3)$$

Отсюда

$$Q'=Q^{-1}, \qquad (4)$$

а поэтому справедливо и равенство

$$QQ'=E. \qquad (5)$$

Таким образом, ввиду (4), *ортогональную матрицу Q можно определить как такую матрицу, для которой транспонированная матрица Q' равна обратной матрице Q^{-1}.* Каждое из равенств (3) и (5) также может быть принято в качестве определения ортогональной матрицы.

Так как столбцы матрицы Q' являются строками матрицы Q, то из (5) вытекает следующее утверждение: *квадратная матрица Q тогда и только тогда будет ортогональной, если сумма квадратов всех элементов любой ее строки равна единице, а сумма произведений соответственных элементов любых двух ее различных строк равна нулю.* Из (3) следует аналогичное утверждение для столбцов матрицы Q.

Переходя в равенстве (3) к определителям, мы получим ввиду того, что $|Q'|=|Q|$, равенство

$$|Q|^2=1.$$

Отсюда следует, что *определитель ортогональной матрицы равен* ± 1. Таким образом, *всякое ортогональное преобразование неизвестных является невырожденным*. Само собой разумеется, что утверждать обратное нельзя; отметим также, что далеко не всякая матрица с определителем, равным ± 1, будет ортогональной.

Матрица, обратная к ортогональной, сама будет ортогональной. Действительно, переходя в (4) к транспонированным матрицам, мы получим:

$$(Q^{-1})' = (Q')' = Q = (Q^{-1})^{-1}.$$

С другой стороны, *произведение ортогональных матриц само ортогонально*. Действительно, если матрицы Q и R ортогональные, то, используя (4), а также равенство (6) из § 26 и аналогичное равенство, справедливое для обратной матрицы, мы получим:

$$(QR)' = R'Q' = R^{-1}Q^{-1} = (QR)^{-1}.$$

В § 37 будет использовано следующее утверждение:

Матрица перехода от ортонормированной базы евклидова пространства к любой другой его ортонормированной базе является ортогональной.

Пусть, в самом деле, в пространстве E_n заданы две ортонормированные базы $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ с матрицей перехода $Q = (q_{ij})$,

$$e' = Qe.$$

Так как база e ортонормированная, то скалярное произведение любых двух векторов, в частности любых двух векторов из базы e', равно сумме произведений соответственных координат этих векторов в базе e. Так как, однако, и база e' ортонормированная, то скалярный квадрат каждого вектора из e' равен единице, а скалярное произведение любых двух разных векторов из e' равно нулю. Отсюда для строк координат векторов базы e' в базе e, т. е. для строк матрицы Q, вытекают те утверждения, которые, как выведено выше из равенства (5), характерны для ортогональной матрицы.

Ортогональные преобразования евклидова пространства. Сейчас уместно изучить один интересный специальный тип линейных преобразований евклидовых пространств, хотя преобразования этого типа и не будут у нас дальше использоваться.

Линейное преобразование φ евклидова пространства E_n называется *ортогональным преобразованием этого евклидова пространства*, если оно сохраняет скалярный квадрат всякого вектора, т. е. для любого вектора a

$$(a\varphi,\ a\varphi) = (a,\ a). \tag{6}$$

Отсюда выводится следующее более общее утверждение, которое, понятно, также может быть принято в качестве определения ортогонального преобразования:

Ортогональное преобразование φ *евклидова пространства сохраняет скалярное произведение любых двух векторов* a, b,

$$(a\varphi,\ b\varphi) = (a,\ b). \tag{7}$$

Действительно, ввиду (6)

$$((a+b)\varphi,\ (a+b)\varphi) = (a+b,\ a+b).$$

Однако

$$((a+b)\varphi,\ (a+b)\varphi) = (a\varphi + b\varphi,\ a\varphi + b\varphi) =$$
$$= (a\varphi,\ a\varphi) + (a\varphi,\ b\varphi) + (b\varphi,\ a\varphi) + (b\varphi,\ b\varphi),$$
$$(a+b,\ a+b) = (a,\ a) + (a,\ b) + (b,\ a) + (b,\ b).$$

Отсюда, используя (6) как для a, так и для b, и учитывая коммутативность скалярного умножения, получаем

$$2(a\varphi,\ b\varphi) = 2(a,\ b),$$

а поэтому имеет место и (7).

При ортогональном преобразовании евклидова пространства образы всех векторов любой ортонормированной базы сами составляют ортонормированную базу. Обратно, если линейное преобразование евклидова пространства переводит хотя бы одну ортонормированную базу снова в ортонормированную базу, то это преобразование ортогонально.

В самом деле, пусть φ — ортогональное преобразование пространства E_n, а $e_1, e_2, \ldots, e_n$ — произвольная ортонормированная база этого пространства. Ввиду (7) из равенств

$$(e_i,\ e_i) = 1, \quad i = 1,\ 2,\ \ldots,\ n,$$
$$(e_i,\ e_j) = 0 \text{ при } i \neq j$$

вытекают равенства

$$(e_i\varphi,\ e_i\varphi) = 1,\ i = 1,\ 2,\ \ldots,\ n,$$
$$(e_i\varphi,\ e_j\varphi) = 0 \text{ при } i \neq j,$$

т. е. система векторов $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ оказывается ортогональной и нормированной, а поэтому она будет ортонормированной базой пространства E_n.

Обратно, пусть линейное преобразование φ пространства E_n переводит ортонормированную базу $e_1, e_2, \ldots, e_n$ снова в ортонормированную базу, т. е. система векторов $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ является ортонормированной базой пространства E_n. Если

$$a = \sum_{i=1}^{n} \alpha_i e_i$$

— произвольный вектор пространства E_n, то

$$a\varphi = \sum_{i=1}^{n} \alpha_i (e_i\varphi),$$

т. е. вектор $a\varphi$ имеет в базе $e\varphi$ те же координаты, что и вектор a в базе e. Эти обе базы являются, однако, ортонормированными, а поэтому скалярный квадрат любого вектора равен сумме квадратов его координат в любой из этих баз. Таким образом,

$$(a,\ a) = (a\varphi,\ a\varphi) = \sum_{i=1}^{n} \alpha_i^2,$$

т. е. равенство (6) действительно выполняется.

Ортогональное преобразование евклидова пространства в любой ортонормированной базе задается ортогональной матрицей. Обратно, если линейное преобразование евклидова пространства хотя бы в одной ортонормированной базе задается ортогональной матрицей, то это преобразование ортогонально.

Действительно, если преобразование φ ортогональное, а база $e_1, e_2, \ldots, e_n$ ортонормированная, то и система векторов $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$ будет ортонормированной базой. Матрица A преобразования φ в базе e,

$$e\varphi = Ae, \tag{8}$$

будет, следовательно, матрицей перехода от ортонормированной базы e к ортонормированной базе $e\varphi$, т. е., как доказано выше, будет ортогональной.

Обратно, пусть линейное преобразование φ задается в ортонормированной базе $e_1, e_2, \ldots, e_n$ ортогональной матрицей A; имеет место, следовательно, равенство (8). Так как база e ортонормированная, то скалярное произведение любых векторов, в частности любых векторов из системы $e_1\varphi, e_2\varphi, \ldots, e_n\varphi$, равно сумме произведений соответственных координат этих векторов в базе e. Поэтому, так как матрица A ортогональна,

$$(e_i\varphi,\ e_i\varphi) = 1,\quad i = 1, 2, \ldots, n,$$
$$(e_i\varphi,\ e_j\varphi) = 0 \ \text{при}\ i \neq j,$$

т. е. система $e\varphi$ сама оказывается ортонормированной базой пространства E_n. Отсюда вытекает ортогональность преобразования φ.

Как читатель знает из курса аналитической геометрии, среди всех аффинных преобразований плоскости, оставляющих на месте начало координат, вращения (соединенные, быть может, с зеркальными отражениями) являются единственными, сохраняющими

скалярное произведение векторов. Таким образом, ортогональные преобразования n-мерного евклидова пространства можно рассматривать как «вращения» этого пространства.

К числу ортогональных преобразований евклидова пространства принадлежит, очевидно, тождественное преобразование. С другой стороны, установленная нами связь между ортогональными преобразованиями и ортогональными матрицами, а также изложенная в § 31 связь между операциями над линейными преобразованиями и над матрицами позволяют из известных свойств ортогональных матриц вывести следующие свойства ортогональных преобразований евклидова пространства, легко проверяемые и непосредственно:

Всякое ортогональное преобразование является невырожденным и его обратное преобразование также ортогонально.

Произведение любых ортогональных преобразований ортогонально.

§ 36. Симметрические преобразования

Линейное преобразование φ n-мерного евклидова пространства называется ***симметрическим*** (или ***самосопряженным***), если для любых векторов a, b этого пространства имеет место равенство

$$(a\varphi,\ b)=(a,\ b\varphi), \tag{1}$$

т. е. символ симметрического преобразования можно при скалярном умножении переносить с одного множителя на другой.

Примерами симметрических преобразований служат, очевидно, тождественное преобразование ε и нулевое преобразование ω. Более общим примером является линейное преобразование, при котором всякий вектор умножается на фиксированное число α,

$$a\varphi=\alpha a.$$

Действительно, в этом случае

$$(a\varphi,\ b)=(\alpha a,\ b)=\alpha(a,\ b)=(a,\ \alpha b)=(a,\ b\varphi).$$

Роль симметрических преобразований весьма велика и нам необходимо изучить их достаточно детально.

Симметрическое преобразование евклидова пространства в любой ортонормированной базе задается симметрической матрицей. Обратно, если линейное преобразование евклидова пространства хотя бы в одной ортонормированной базе задается симметрической матрицей, то это преобразование симметрическое.

Действительно, пусть симметрическое преобразование φ задается в ортонормированной базе $e_1, e_2, \ldots, e_n$ матрицей $A=(\alpha_{ij})$. Учитывая, что в ортонормированной базе скалярное произведение двух

векторов равно сумме произведений соответственных координат этих векторов, мы получаем:

$$(e_i\varphi,\ e_j)=\left(\sum_{k=1}^{n}\alpha_{ik}e_k,\ e_j\right)=\alpha_{ij},$$

$$(e_i,\ e_j\varphi)=\left(e_i,\ \sum_{k=1}^{n}\alpha_{jk}e_k\right)=\alpha_{ji},$$

т. е., ввиду (1),

$$\alpha_{ij}=\alpha_{ji}$$

для всех i и j. Матрица A оказалась, таким образом, симметрической.

Обратно, пусть линейное преобразование φ задается в ортонормированной базе $e_1,\ e_2,\ \ldots,\ e_n$ симметрической матрицей $A=(\alpha_{ij})$,

$$\alpha_{ij}=\alpha_{ji}\quad \text{для всех } i \text{ и } j. \tag{2}$$

Если

$$b=\sum_{i=1}^{n}\beta_i e_i,\qquad c=\sum_{j=1}^{n}\gamma_j e_j$$

— любые векторы пространства, то

$$b\varphi=\sum_{i=1}^{n}\beta_i\,(e_i\varphi)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}\beta_i\alpha_{ij}\right)e_j,$$

$$c\varphi=\sum_{j=1}^{n}\gamma_j\,(e_j\varphi)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\gamma_j\alpha_{ji}\right)e_i.$$

Используя ортонормированность базы e, получаем

$$(b\varphi,\ c)=\sum_{j,\,i=1}^{n}\beta_i\alpha_{ij}\gamma_j,$$

$$(b,\ c\varphi)=\sum_{i,\,j=1}^{n}\beta_i\gamma_j\alpha_{ji}.$$

Ввиду (2) правые части последних равенств совпадают, а поэтому

$$(b\varphi,\ c)=(b,\ c\varphi),$$

что и требовалось доказать.

Из полученного результата вытекает следующее свойство симметрических преобразований, легко проверяемое и непосредственно:

Сумма симметрических преобразований, а также произведение симметрического преобразования на число являются симметрическими преобразованиями.

Докажем теперь следующую важную теорему:

Все характеристические корни симметрического преобразования действительны.

Так как характеристические корни любого линейного преобразования совпадают с характеристическими корнями матрицы этого преобразования в любой базе, а симметрическое преобразование задается в ортонормированных базах симметрическими матрицами, то достаточно доказать следующее утверждение:

Все характеристические корни симметрической матрицы действительны.

В самом деле, пусть λ_0 будет характеристический корень (быть может, комплексный) симметрической матрицы $A=(\alpha_{ij})$,

$$|A-\lambda_0 E|=0.$$

Тогда система линейных однородных уравнений с комплексными коэффициентами

$$\sum_{j=1}^{n}\alpha_{ij}x_j=\lambda_0 x_i, \qquad i=1,\ 2,\ \ldots,\ n,$$

имеет равный нулю определитель, т. е. обладает ненулевым решением $\beta_1,\ \beta_2,\ \ldots,\ \beta_n$, вообще говоря, комплексным; таким образом,

$$\sum_{j=1}^{n}\alpha_{ij}\beta_j=\lambda_0\beta_i,\quad i=1,\ 2,\ \ldots,\ n. \tag{3}$$

Умножая обе части каждого i-го из равенств (3) на число $\overline{\beta}_i$, сопряженное с числом β_i, и складывая отдельно левые и правые части всех получающихся равенств, мы приходим к равенству

$$\sum_{i,\ j=1}^{n}\alpha_{ij}\beta_j\overline{\beta}_i=\lambda_0\sum_{i=1}^{n}\beta_i\overline{\beta}_i. \tag{4}$$

Коэффициент при λ_0 в (4) является отличным от нуля действительным числом, будучи суммой неотрицательных действительных чисел, хотя бы одно из которых строго положительно. Действительность числа λ_0 будет поэтому доказана, если мы докажем действительность левой части равенства (4), для чего достаточно показать, что это комплексное число совпадает со своим сопряженным. Здесь впервые будет использована симметричность (действительной) матрицы A.

$$\overline{\sum_{i,\ j=1}^{n}\alpha_{ij}\beta_j\overline{\beta}_i}=\sum_{i,\ j=1}^{n}\overline{\alpha_{ij}\beta_j\overline{\beta}_i}=\sum_{i,\ j=1}^{n}\alpha_{ij}\overline{\beta}_j\beta_i=$$
$$=\sum_{i,\ j=1}^{n}\alpha_{ji}\overline{\beta}_j\beta_i=\sum_{i,\ j=1}^{n}\alpha_{ij}\overline{\beta}_i\beta_j=\sum_{i,\ j=1}^{n}\alpha_{ij}\beta_j\overline{\beta}_i.$$

Заметим, что предпоследнее равенство получено простой переменой обозначений для индексов суммирования: вместо i поставлено j, вместо j поставлено i. Теорема, следовательно, доказана.

Линейное преобразование φ евклидова пространства E_n тогда и только тогда будет симметрическим, если в пространстве E_n существует ортонормированная база, составленная из собственных векторов этого преобразования.

В одну сторону это утверждение почти очевидно: если в E_n существует ортонормированная база $e_1, e_2, \ldots, e_n$, причем

$$e_i\varphi = \lambda_i e_i, \quad i = 1, 2, \ldots, n,$$

то в базе e преобразование φ задается диагональной матрицей

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Диагональная матрица является, однако, симметрической, а поэтому преобразование φ задается в ортонормированной базе e симметрической матрицей, т. е. будет симметрическим.

Основное обратное утверждение теоремы мы будем доказывать индукцией по размерности n пространства E_n. В самом деле, при $n = 1$ всякое линейное преобразование φ пространства E_1 непременно переводит любой вектор в вектор, ему пропорциональный. Отсюда следует, что всякий ненулевой вектор a будет собственным вектором для φ (как, впрочем, следует и то, что всякое линейное преобразование пространства E_1 будет симметрическим). Нормируя вектор a, мы получим искомую ортонормированную базу пространства E_1.

Пусть утверждение теоремы уже доказано для $(n-1)$-мерного евклидова пространства и пусть в пространстве E_n задано симметрическое преобразование φ. Из доказанной выше теоремы вытекает существование для φ действительного характеристического корня λ_0. Это число будет, следовательно, собственным значением для преобразования φ. Если a — собственный вектор преобразования φ, относящийся к этому собственному значению, то и всякий ненулевой вектор, пропорциональный вектору a, будет для φ собственным вектором, относящимся к тому же собственному значению λ_0, так как

$$(\alpha a)\varphi = \alpha(a\varphi) = \alpha(\lambda_0 a) = \lambda_0(\alpha a).$$

В частности, нормируя вектор a, мы получим такой вектор e_1, что

$$e_1\varphi = \lambda_0 e_1,$$
$$(e_1, e_1) = 1.$$

Как доказано в § 34, ненулевой вектор e_1 можно включить в ортогональную базу

$$e_1, e_2', \ldots, e_n', \tag{5}$$

пространства E_n. Те векторы, первая координата которых в базе (5) равна нулю, т. е. векторы вида $\alpha_2 e_2' + \ldots + \alpha_n e_n'$, составляют, очевидно, $(n-1)$-мерное линейное подпространство пространства E_n, которое мы обозначим через L. Это будет даже $(n-1)$-мерное евклидово пространство, так как скалярное произведение, будучи определенным для всех векторов из E_n, определено, в частности, для векторов из L, причем обладает всеми необходимыми свойствами.

Подпространство L состоит из всех тех векторов пространства E_n, которые ортогональны к вектору e_1. Действительно, если

$$a = \alpha_1 e_1 + \alpha_2' e_2' + \ldots + \alpha_n' e_n',$$

то, ввиду ортогональности базы (5) и нормированности вектора e_1,

$$(e_1, a) = \alpha_1 (e_1, e_1) + \alpha_2' (e_1, e_2') + \ldots + \alpha_n' (e_1, e_n') = \alpha_1,$$

т. е. $(e_1, a) = 0$ тогда и только тогда, если $\alpha_1 = 0$.

Если вектор a принадлежит к подпространству L, т. е. $(e_1, a) = 0$, то и вектор $a\varphi$ содержится в L. Действительно, ввиду симметричности преобразования φ,

$$(e_1, a\varphi) = (e_1\varphi, a) = (\lambda_0 e_1, a) = \lambda_0 (e_1, a) = \lambda_0 \cdot 0 = 0,$$

т. е. вектор $a\varphi$ ортогонален к e_1 и поэтому содержится в L. Это свойство подпространства L, называемое его *инвариантностью* относительно преобразования φ, позволяет считать φ, рассматриваемое лишь в применении к векторам из L, линейным преобразованием этого $(n-1)$-мерного евклидова пространства. Оно будет даже симметрическим преобразованием пространства L, так как равенство (1), выполняясь для любых векторов из E_n, будет выполняться, в частности, для векторов, лежащих в L.

В силу индуктивного предположения в пространстве L существует ортонормированная база, состоящая из собственных векторов преобразования φ; обозначим ее через $e_2, \ldots, e_n$. Все эти векторы ортогональны к вектору e_1, а поэтому $e_1, e_2, \ldots, e_n$ будет искомой ортонормированной базой пространства E_n, состоящей из собственных векторов преобразования φ. Теорема доказана.

§ 37. Приведение квадратичной формы к главным осям. Пары форм

Применим последнюю теорему предшествующего параграфа к доказательству следующей матричной теоремы:

Для всякой симметрической матрицы A можно найти такую ортогональную матрицу Q, которая приводит матрицу A к диагональному виду, т. е. матрица $Q^{-1}AQ$, полученная трансформированием матрицы A матрицей Q, будет диагональной.

В самом деле, пусть дана симметрическая матрица A порядка n. Если $e_1, e_2, \ldots, e_n$ — некоторая ортонормированная база n-мерного евклидова пространства E_n, то матрица A задает в этой базе симметрическое преобразование φ. Как доказано, в E_n существует ортонормированная база $f_1, f_2, \ldots, f_n$, составленная из собственных векторов преобразования φ; в этой базе φ задается диагональной матрицей B (см. § 33). Тогда, по § 31,

$$B = Q^{-1}AQ, \tag{1}$$

где Q — матрица перехода от базы f к базе e,

$$e = Qf. \tag{2}$$

Эта матрица, как матрица перехода от одной ортонормированной базы к другой такой же базе, будет ортогональной — см. § 35. Теорема доказана.

Так как для ортогональной матрицы Q ее обратная матрица равна транспонированной, $Q^{-1} = Q'$, то равенство (1) можно переписать в виде

$$B = Q'AQ.$$

Из § 26 известно, однако, что именно так преобразуется симметрическая матрица A квадратичной формы, подвергнутой линейному преобразованию неизвестных с матрицей Q. Учитывая же, что линейное преобразование неизвестных с ортогональной матрицей является ортогональным преобразованием (см. § 35) и что диагональную матрицу имеет квадратичная форма, приведенная к каноническому виду, мы на основании предшествующей теоремы получаем следующую теорему о приведении действительной квадратичной формы к главным осям:

Всякая действительная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ некоторым ортогональным преобразованием неизвестных может быть приведена к каноническому виду.

Хотя может существовать много различных ортогональных преобразований неизвестных, приводящих данную квадратичную форму к каноническому виду, однако сам этот канонический вид по существу определяется однозначно:

Каково бы ни было ортогональное преобразование, приводящее к каноническому виду квадратичную форму $f(x_1, x_2, \ldots, x_n)$ с матрицей A, коэффициентами этого канонического вида будут характеристические корни матрицы A, взятые с их кратностями.

Пусть, в самом деле, форма f некоторым ортогональным преобразованием приведена к каноническому виду

$$f(x_1, x_2, \ldots, x_n) = \mu_1 y_1^2 + \mu_2 y_2^2 + \ldots + \mu_n y_n^2.$$

Это ортогональное преобразование оставляет инвариантной сумму квадратов неизвестных, а поэтому, если λ — новое неизвестное, то

$$f(x_1, x_2, \ldots, x_n) - \lambda \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} \mu_i y_i^2 - \lambda \sum_{i=1}^{n} y_i^2.$$

Переходя к определителям этих квадратичных форм и учитывая, что после выполнения линейного преобразования определитель квадратичной формы умножается на квадрат определителя преобразования (см. § 28), а квадрат определителя ортогонального преобразования равен единице (см. § 35), мы приходим к равенству

$$|A - \lambda E| = \begin{vmatrix} \mu_1 - \lambda & 0 & \ldots & 0 \\ 0 & \mu_2 - \lambda & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_n - \lambda \end{vmatrix} = \prod_{i=1}^{n} (\mu_i - \lambda),$$

из которого вытекает утверждение теоремы.

Этому результату можно придать также матричную формулировку: *Какова бы ни была ортогональная матрица, приводящая к диагональному виду симметрическую матрицу A, на главной диагонали полученной диагональной матрицы будут стоять характеристические корни матрицы A, взятые с их кратностями.*

Практическое разыскание ортогонального преобразования, приводящего квадратичную форму к главным осям. В некоторых задачах необходимо знать не только тот канонический вид, к которому приводится действительная квадратичная форма ортогональным преобразованием, но и само ортогональное преобразование, осуществляющее это приведение. Было бы затруднительно разыскивать это преобразование, используя доказательство теоремы о приведении к главным осям, и мы хотим указать иной путь. Именно, нужно лишь научиться находить ортогональную матрицу Q, приводящую данную симметрическую матрицу A к диагональному виду, или, что то же самое, находить ее обратную матрицу Q^{-1}. Ввиду (2) это будет матрица перехода от базы e к базе f, т. е. ее строки являются координатными строками (в базе e) ортонормированной системы из n собственных векторов симметрического преобразования φ, определяемого матрицей A в базе e. Остается найти такую систему собственных векторов.

Пусть λ_0 — любой характеристический корень матрицы A и пусть его кратность равна k_0. Из § 33 мы знаем, что совокупность координатных строк всех собственных векторов преобразования φ, относящихся к собственному значению λ_0, совпадает с совокупностью ненулевых решений системы линейных однородных уравнений

$$(A - \lambda_0 E) X = 0; \tag{3}$$

симметричность матрицы A позволяет написать здесь A вместо A' Из доказанных выше теорем существования ортогональной матрицы,

приводящей симметрическую матрицу A к диагональному виду, и единственности этого диагонального вида вытекает, что для системы (3) во всяком случае можно найти k_0 линейно независимых решений. Такую систему решений ищем методами, известными из § 12, а затем ортогонализируем и нормируем полученную систему в соответствии с § 34.

Беря в качестве λ_0 поочередно все различные характеристические корни симметрической матрицы A и учитывая, что сумма кратностей этих корней равна n, мы получим систему из n собственных векторов преобразования φ, заданных их координатами в базе e. Для доказательства того, что это будет искомая ортонормированная система собственных векторов, остается доказать следующую лемму:

Собственные векторы симметрического преобразования φ, относящиеся к различным собственным значениям, между собой ортогональны.

Пусть, в самом деле,

$$b\varphi=\lambda_1 b, \qquad c\varphi=\lambda_2 c,$$

причем $\lambda_1 \neq \lambda_2$. Так как

$$(b\varphi,\ c)=(\lambda_1 b,\ c)=\lambda_1 (b,\ c),$$
$$(b,\ c\varphi)=(b,\ \lambda_2 c)=\lambda_2 (b,\ c),$$

то из

$$(b\varphi,\ c)=(b,\ c\varphi)$$

следует

$$\lambda_1 (b,\ c)=\lambda_2 (b,\ c)$$

или, ввиду $\lambda_1 \neq \lambda_2$,

$$(b,\ c)=0,$$

что и требовалось доказать.

Пример. Привести к главным осям квадратичную форму

$$f(x_1,\ x_2,\ x_3,\ x_4)=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4.$$

Матрица A этой формы имеет вид

$$A=\begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Найдем ее характеристический многочлен:

$$|A-\lambda E|=\begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix}=(\lambda-1)^3(\lambda+3).$$

Таким образом, матрица A имеет трехкратный характеристический корень 1 и простой характеристический корень -3. Мы уже можем, следовательно, написать тот канонический вид, к которому форма f приводится ортогональным преобразованием:

$$f=y_1^2+y_2^2+y_3^2-3y_4^2.$$

Найдем ортогональное преобразование, осуществляющее это приведение. Система линейных однородных уравнений (3) при $\lambda_0=1$ принимает вид

$$\begin{cases} -x_1+x_2+x_3-x_4=0, \\ x_1-x_2-x_3+x_4=0, \\ x_1-x_2-x_3+x_4=0, \\ -x_1+x_2+x_3-x_4=0. \end{cases}$$

Ранг этой системы равен 1, и поэтому для нее можно найти три линейно независимых решения. Ими будут, например, векторы

$$b_1=(1,\ 1,\ 0,\ 0),$$
$$b_2=(1,\ 0,\ 1,\ 0),$$
$$b_3=(-1,\ 0,\ 0,\ 1).$$

Ортогонализируя эту систему векторов, мы получим систему векторов

$$c_1=b_1=(1,\ 1,\ 0,\ 0),$$
$$c_2=-\frac{1}{2}c_1+b_2=\left(\frac{1}{2},-\frac{1}{2},\ 1,\ 0\right),$$
$$c_3=\frac{1}{2}c_1+\frac{1}{3}c_2+b_3=\left(-\frac{1}{3},\frac{1}{3},\frac{1}{3},1\right).$$

С другой стороны, система линейных однородных уравнений (3) принимает при $\lambda_0=-3$ вид

$$\begin{cases} 3x_1+x_2+x_3-x_4=0, \\ x_1+3x_2-x_3+x_4=0, \\ x_1-x_2+3x_3+x_4=0, \\ -x_1+x_2+x_3+3x_4=0. \end{cases}$$

Ранг этой системы равен 3. Ее ненулевым решением служит вектор

$$c_4=(1,\ -1,\ -1,\ 1).$$

Система векторов c_1, c_2, c_3, c_4 ортогональная. Нормируя ее, мы придем к ортонормированной системе векторов

$$c_1'=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},\ 0,\ 0\right),$$
$$c_2'=\left(\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\sqrt{\frac{2}{3}},0\right),$$
$$c_3'=\left(-\frac{1}{2\sqrt{3}},\frac{1}{2\sqrt{3}},\frac{1}{2\sqrt{3}},\frac{\sqrt{3}}{2}\right),$$
$$c_4'=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right).$$

Таким образом, форма f приводится к главным осям ортогональным преобразованием

$$
\begin{aligned}
y_1 &= \frac{1}{\sqrt{2}} x_1 + \frac{1}{\sqrt{2}} x_2, \\
y_2 &= \frac{1}{\sqrt{6}} x_1 - \frac{1}{\sqrt{6}} x_2 + \sqrt{\frac{2}{3}} x_3, \\
y_3 &= -\frac{1}{2\sqrt{3}} x_1 + \frac{1}{2\sqrt{3}} x_2 + \frac{1}{2\sqrt{3}} x_3 + \frac{\sqrt{3}}{2} x_4, \\
y_4 &= \frac{1}{2} x_1 - \frac{1}{2} x_2 - \frac{1}{2} x_3 + \frac{1}{2} x_4.
\end{aligned}
$$

Следует отметить, что выбор системы линейно независимых собственных векторов, относящихся к кратиому собственному значению, является весьма неоднозначным, а поэтому существует много различных ортогональных преобразований, приводящих форму f к каноническому виду. Мы нашли лишь одно из них.

Пары форм. Пусть дана пара действительных квадратичных форм от n неизвестных, $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$. Существует ли такое невырожденное линейное преобразование неизвестных $x_1, x_2, \ldots, x_n$, которое одновременно приводило бы обе эти формы к каноническому виду?

В общем случае ответ будет отрицательным. Рассмотрим, например, пару форм

$$f(x_1, x_2) = x_1^2, \qquad g(x_1, x_2) = x_1 x_2.$$

Пусть существует невырожденное линейное преобразование

$$\left.\begin{aligned} x_1 &= c_{11} y_1 + c_{12} y_2, \\ x_2 &= c_{21} y_1 + c_{22} y_2, \end{aligned}\right\} \tag{4}$$

приводящее обе эти формы к каноническому виду. Для того чтобы форма f могла быть приведена преобразованием (4) к каноническому виду, один из коэффициентов c_{11}, c_{12} должен быть равен нулю, иначе вошел бы член $2c_{11}c_{12}y_1y_2$. Меняя, если нужно, нумерацию неизвестных y_1, y_2, можно положить, что $c_{12}=0$ и поэтому $c_{11} \neq 0$. Мы получим теперь, однако, что

$$g(x_1, x_2) = c_{11}y_1(c_{21}y_1 + c_{22}y_2) = c_{11}c_{21}y_1^2 + c_{11}c_{22}y_1y_2.$$

Так как форма g также должна была перейти в канонический вид, то $c_{11}c_{22}=0$, т. е. $c_{22}=0$, что вместе с $c_{12}=0$ противоречит невырожденности линейного преобразования (4).

Ситуация будет иной, если мы положим, что хотя бы одна из наших форм, например $g(x_1, x_2, \ldots, x_n)$, является *положительно определенной*[1]. Именно, справедлива теорема:

---

[1]) Это условие не является, конечно, необходимым; так, формы $x_1^2 + x_2^2 - x_3^2$ и $x_1^2 - x_2^2 - x_3^2$ обе уже имеют канонический вид, хотя среди них нет положительно определенных.

Если f и g — пара действительных квадратичных форм от n неизвестных, причем вторая из них положительно определенная, то существует невырожденное линейное преобразование, одновременно приводящее форму g к нормальному виду, а форму f к каноническому виду.

Для доказательства выполним сначала невырожденное линейное преобразование неизвестных $x_1, x_2, \ldots, x_n$,

$$X = TY,$$

приводящее положительно определенную форму g к нормальному виду,

$$g(x_1, x_2, \ldots, x_n) = y_1^2 + y_2^2 + \ldots + y_n^2.$$

Форма f перейдет при этом в некоторую форму φ от новых неизвестных,

$$f(x_1, x_2, \ldots, x_n) = \varphi(y_1, y_2, \ldots, y_n).$$

Совершим теперь ортогональное преобразование неизвестных $y_1, y_2, \ldots, y_n$,

$$Y = QZ,$$

приводящее форму φ к главным осям,

$$\varphi(y_1, y_2, \ldots, y_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

Это преобразование (см. определение в § 35) переводит сумму квадратов неизвестных $y_1, y_2, \ldots, y_n$ в сумму квадратов неизвестных $z_1, z_2, \ldots, z_n$. В результате мы получаем

$$f(x_1, x_2, \ldots, x_n) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2,$$
$$g(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \ldots + z_n^2,$$

т. е. линейное преобразование

$$X = (TQ) Z$$

является искомым.

ГЛАВА ДЕВЯТАЯ

ВЫЧИСЛЕНИЕ КОРНЕЙ МНОГОЧЛЕНОВ

§ 38*. Уравнения второй, третьей и четвертой степени

Основная теорема, доказанная в § 23, устанавливает для любого многочлена n-й степени с числовыми коэффициентами существование n комплексных корней. Ее доказательства (как приведенное выше, так и любые другие из ныне известных) не дают, однако, никаких методов для практического разыскания этих корней, являясь чистыми «доказательствами существования». Поиски таких методов начались, естественно, с попыток вывода формул, аналогичных формуле для решения квадратного уравнения, известной читателю для случая действительных коэффициентов из школьного курса алгебры. Мы покажем сейчас, что эта формула остается справедливой и для квадратных уравнений с комплексными коэффициентами и что аналогичные формулы, хотя и много более громоздкие, могут быть выведены для уравнений третьей и четвертой степени.

Квадратные уравнения. Пусть дано квадратное уравнение

$$x^2+px+q=0$$

с любыми комплексными коэффициентами; старший коэффициент без ограничения общности можно считать равным единице. Это уравнение можно переписать в виде

$$\left(x+\frac{p}{2}\right)^2+\left(q-\frac{p^2}{4}\right)=0.$$

Как мы знаем, из комплексного числа $\frac{p^2}{4}-q$ можно извлечь квадратный корень, не выходя за пределы системы комплексных чисел. Два значения этого корня, отличающихся друг от друга лишь знаком, мы запишем в виде $\pm\sqrt{\frac{p^2}{4}-q}$. Поэтому

$$x+\frac{p}{2}=\pm\sqrt{\frac{p^2}{4}-q},$$

т. е. корни заданного уравнения можно находить по обычной формуле

$$x=-\frac{p}{2}\pm\sqrt{\frac{p^2}{4}-q}.$$

Пример. Решить уравнение

$$x^2-3x+(3-i)=0.$$

Применяя выведенную формулу, получаем:

$$x=\frac{3}{2}\pm\sqrt{\frac{9}{4}-(3-i)}=\frac{3}{2}\pm\frac{1}{2}\sqrt{-3+4i}.$$

При помощи методов § 19 мы находим:

$$\sqrt{-3+4i}=\pm(1+2i),$$

а поэтому

$$x_1=2+i,\quad x_2=1-i.$$

Кубичные уравнения. В отличие от случая квадратных уравнений, до сих пор у нас не было метода для решения кубичных уравнений даже в случае действительных коэффициентов. Сейчас мы выведем для кубичных уравнений формулу, аналогичную формуле для квадратных уравнений, причем сразу допустим, что коэффициенты являются любыми комплексными числами.

Пусть дано кубичное уравнение

$$y^3+ay^2+by+c=0 \tag{1}$$

с любыми комплексными коэффициентами. Заменяя в уравнении (1) неизвестное y новым неизвестным x, связанным с y равенством

$$y=x-\frac{a}{3}, \tag{2}$$

мы получим уравнение относительно неизвестного x, не содержащее, как легко проверить, квадрата этого неизвестного, т. е. уравнение вида

$$x^3+px+q=0. \tag{3}$$

Если будут найдены корни уравнения (3), то, ввиду (2), мы получим и корни заданного уравнения (1). Нам остается, следовательно, научиться решать «неполное» кубичное уравнение (3) с любыми комплексными коэффициентами.

Уравнение (3) обладает по основной теореме тремя комплексными корнями. Пусть x_0 будет любой из этих корней. Введем вспомогательное неизвестное u и рассмотрим многочлен

$$f(u)=u^2-x_0u-\frac{p}{3}.$$

Его коэффициенты — комплексные числа, и поэтому он обладает двумя комплексными корнями α и β, причем, по формулам Вьета,

$$\alpha+\beta=x_0, \tag{4}$$

$$\alpha\beta=-\frac{p}{3}. \tag{5}$$

Подставляя в (3) выражение (4) корня x_0, мы получим:

$$(\alpha+\beta)^3+p(\alpha+\beta)+q=0$$

или

$$\alpha^3+\beta^3+(3\alpha\beta+p)(\alpha+\beta)+q=0.$$

Однако из (5) следует $3\alpha\beta+p=0$, и поэтому мы получаем:

$$\alpha^3+\beta^3=-q. \tag{6}$$

С другой стороны, из (5) вытекает

$$\alpha^3\beta^3=-\frac{p^3}{27}. \tag{7}$$

Равенства (6) и (7) показывают, что числа α^3 и β^3 служат корнями квадратного уравнения

$$z^2+qz-\frac{p^3}{27}=0 \tag{8}$$

с комплексными коэффициентами.

Решая уравнение (8), мы получим:

$$z=-\frac{q}{2}\pm\sqrt{\frac{q^2}{4}+\frac{p^3}{27}},$$

откуда [1])

$$\alpha=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}},\quad \beta=\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}. \tag{9}$$

Мы приходим к следующей *формуле Кардано*, выражающей корни уравнения (3) через его коэффициенты при помощи квадратных и кубичных радикалов:

$$x_0=\alpha+\beta=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}+\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}.$$

Так как кубичный радикал имеет в поле комплексных чисел три значения, то формулы (9) дают три значения для α и три для β. Нельзя, однако, применяя формулу Кардано, комбинировать любое значение радикала α с любым значением радикала β: для данного значения α следует брать лишь то из трех значений β, которое удовлетворяет условию (5).

---

[1]) Безразлично, какой из корней уравнения (8) принять за α^3 и какой — за β^3, так как α и β в равенстве (6) и (7), а также в выражение (4) для x_0 входят симметричным образом.

Пусть α_1 будет любое из трех значений радикала α. Тогда два других можно получить, как доказано в § 19, умножением α_1 на кубичные корни ε и ε^2 из единицы:

$$\alpha_2 = \alpha_1\varepsilon, \quad \alpha_3 = \alpha_1\varepsilon^2.$$

Обозначим через β_1 то из трех значений радикала β, которое соответствует значению α_1 радикала α на основании (5), т. е. $\alpha_1\beta_1 = -\frac{p}{3}$. Два других значения β будут

$$\beta_2 = \beta_1\varepsilon, \quad \beta_3 = \beta_1\varepsilon^2.$$

Так как, ввиду $\varepsilon^3 = 1$,

$$\alpha_2\beta_3 = \alpha_1\varepsilon\cdot\beta_1\varepsilon^2 = \alpha_1\beta_1\varepsilon^3 = \alpha_1\beta_1 = -\frac{p}{3},$$

то значению α_2 радикала α соответствует значение β_3 радикала β; аналогично значению α_3 соответствует значение β_2. Таким образом, все три корня уравнения (3) могут быть записаны следующим образом:

$$\left.\begin{aligned} x_1 &= \alpha_1 + \beta_1, \\ x_2 &= \alpha_2 + \beta_3 = \alpha_1\varepsilon + \beta_1\varepsilon^2, \\ x_3 &= \alpha_3 + \beta_2 = \alpha_1\varepsilon^2 + \beta_1\varepsilon. \end{aligned}\right\} \qquad (10)$$

Кубичные уравнения с действительными коэффициентами. Посмотрим, что можно сказать о корнях неполного кубичного уравнения

$$x^3 + px + q = 0, \qquad (11)$$

если его коэффициенты действительны. Оказывается, что в этом случае основную роль играет знак выражения $\frac{q^2}{4} + \frac{p^3}{27}$, стоящего в формуле Кардано под знаком квадратного корня. Заметим, что знак этого выражения противоположен знаку выражения

$$D = -4p^3 - 27q^2 = -108\left(\frac{q^2}{4} + \frac{p^3}{27}\right),$$

называемого *дискриминантом* уравнения (11) (ср. ниже, § 54); в дальнейших формулировках будет использоваться знак дискриминанта.

1) Пусть $D < 0$. В этом случае в формуле Кардано под знаком каждого из квадратных радикалов стоит положительное число, а поэтому под знаком каждого из кубичных радикалов оказываются действительные числа. Однако кубичный корень из действительного числа имеет одно действительное и два сопряженных комплексных значения. Пусть α_1 будет действительное значение радикала α; тогда значение β_1 радикала β, соответствующее α_1 на основании формулы (5), также будет действительным ввиду действительности числа p. Таким образом, корень $x_1 = \alpha_1 + \beta_1$ уравнения (11) оказы-

вается действительным. Два других корня мы найдем, заменяя в формулах (10) настоящего параграфа корни из единицы $\varepsilon = \varepsilon_1$ и $\varepsilon^2 = \varepsilon_2$ их выражениями (7) из § 19:

$$x_2 = \alpha_1\varepsilon + \beta_1\varepsilon^2 = \alpha_1\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) + \beta_1\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = -\frac{\alpha_1+\beta_1}{2} + i\sqrt{3}\frac{\alpha_1-\beta_1}{2},$$

$$x_3 = \alpha_1\varepsilon^2 + \beta_1\varepsilon = \alpha_1\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) + \beta_1\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = -\frac{\alpha_1+\beta_1}{2} - i\sqrt{3}\frac{\alpha_1-\beta_1}{2};$$

эти два корня оказываются ввиду действительности чисел α_1 и β_1 сопряженными комплексными числами, причем коэффициент при мнимой части отличен от нуля, так как $\alpha_1 \neq \beta_1$, — эти числа являются значениями различных кубичных радикалов.

Таким образом, *если* $D < 0$, *то уравнение* (11) *имеет один действительный и два сопряженных комплексных корня.*

2) Пусть $D = 0$. В этом случае

$$\alpha = \sqrt[3]{-\frac{q}{2}}, \quad \beta = \sqrt[3]{-\frac{q}{2}}.$$

Пусть α_1 будет действительное значение радикала α; тогда β_1 также будет, ввиду (5), действительным числом, причем $\alpha_1 = \beta_1$. Заменяя в формулах (10) β_1 через α_1 и используя очевидное равенство $\varepsilon + \varepsilon^2 = -1$, мы получим:

$$x_1 = 2\alpha_1, \quad x_2 = \alpha_1(\varepsilon + \varepsilon^2) = -\alpha_1, \quad x_3 = \alpha_1(\varepsilon^2 + \varepsilon) = -\alpha_1.$$

Таким образом, *если* $D = 0$, *то все корни уравнения* (11) *действительны, причем два из них равны между собой.*

3) Пусть, наконец, $D > 0$. В этом случае в формуле Кардано под знаком квадратного корня стоит отрицательное действительное число, а поэтому под знаками кубичных радикалов стоят сопряженные комплексные числа. Таким образом, все значения радикалов α и β будут теперь комплексными числами. Среди корней уравнения (11) должен, однако, содержаться хотя бы один действительный. Пусть это будет корень

$$x_1 = \alpha_0 + \beta_0.$$

Так как действительны и сумма чисел α_0 и β_0, и их произведение, равное $-\frac{p}{3}$, то числа α_0 и β_0 сопряжены между собой как корни квадратного уравнения с действительными коэффициентами. Но тогда

сопряжены между собой и числа $\alpha_0\varepsilon$ и $\beta_0\varepsilon^2$, а также числа $\alpha_0\varepsilon^2$ и $\beta_0\varepsilon$, откуда следует, что корни уравнения (11)

$$x_2=\alpha_0\varepsilon+\beta_0\varepsilon^2, \quad x_3=\alpha_0\varepsilon^2+\beta_0\varepsilon$$

также будут действительными числами.

Мы получили, что все три корня уравнения (11) действительны, причем легко показать, что среди них нет равных. В самом деле, в противном случае выбор корня x_1 можно было бы осуществить так, чтобы имело место равенство $x_2=x_3$, откуда

$$\alpha_0(\varepsilon-\varepsilon^2)=\beta_0(\varepsilon-\varepsilon^2),$$

т. е. $\alpha_0=\beta_0$, что явно невозможно.

Таким образом, *если* $D>0$, *то уравнение* (11) *имеет три различных действительных корня.*

Рассмотренный сейчас последний случай показывает, что практическое значение формулы Кардано весьма невелико. В самом деле, хотя при $D>0$ все корни уравнения (11) с действительными коэффициентами являются действительными числами, однако разыскание их по формуле Кардано требует извлечения кубичных корней из комплексных чисел, что мы умеем делать лишь переходом к тригонометрической форме этих чисел. Поэтому запись корней с помощью радикалов теряет практическое значение. При помощи методов, выходящих за рамки нашей книги, можно было бы доказать, что в рассматриваемом случае корни уравнения (11) вообще никаким способом не могут быть выражены через коэффициенты при помощи радикалов с действительными подкоренными выражениями. Этот случай решения уравнения (11) называется *неприводимым* (не смешивать с неприводимостью многочленов!).

Примеры. 1. Решить уравнение

$$y^3+3y^2-3y-14=0.$$

Подстановка $y=x-1$ приводит это уравнение к виду

$$x^3-6x-9=0. \tag{12}$$

Здесь $p=-6$, $q=-9$, поэтому

$$\frac{q^2}{4}+\frac{p^3}{27}=\frac{49}{4}>0,$$

т. е. уравнение (12) имеет один действительный и два сопряженных комплексных корня. По (9) $\alpha=\sqrt[3]{\frac{9}{2}+\frac{7}{2}}=\sqrt[3]{8}$, $\beta=\sqrt[3]{\frac{9}{2}-\frac{7}{2}}=\sqrt[3]{1}$. Поэтому $\alpha_1=2$, $\beta_1=1$; т. е. $x_1=3$. Два других корня найдем по формулам (10): $x_2=-\frac{3}{2}+i\frac{\sqrt{3}}{2}$, $x_3=-\frac{3}{2}-i\frac{\sqrt{3}}{2}$.

Отсюда следует, что корнями заданного уравнения служат числа

$$y_1=2,\ y_2=-\frac{5}{2}+i\frac{\sqrt{3}}{2},\quad y_3=-\frac{5}{2}-i\frac{\sqrt{3}}{2}.$$

2. Решить уравнение

$$x^3-12x+16=0.$$

Здесь $p=-12$, $q=16$, поэтому

$$\frac{q^2}{4}+\frac{p^3}{27}=0.$$

Отсюда следует: $\alpha=\sqrt[3]{-8}$, т. е. $\alpha_1=-2$. Поэтому

$$x_1=-4,\quad x_2=x_3=2.$$

3. Решить уравнение

$$x^3-19x+30=0.$$

Здесь $p=-19$, $q=30$, поэтому

$$\frac{q^2}{4}+\frac{p^3}{27}=-\frac{784}{27}<0.$$

Таким образом, если оставаться в области действительных чисел, формула Кардано к этому уравнению неприменима, хотя его корнями являются действительные числа 2, 3 и -5.

Уравнения четвертой степени. Решение уравнения четвертой степени

$$y^4+ay^3+by^2+cy+d=0 \tag{13}$$

с произвольными комплексными коэффициентами сводится к решению некоторого вспомогательного кубичного уравнения. Достигается это следующим методом, принадлежащим Феррари.

Предварительно уравнение (13) подстановкой $y=x-\frac{a}{4}$ приводится к виду

$$x^4+px^2+qx+r=0. \tag{14}$$

Затем левая часть этого уравнения следующим образом тождественно преобразуется при помощи вспомогательного параметра α:

$$x^4+px^2+qx+r=\left(x^2+\frac{p}{2}+\alpha\right)^2+qx+r-\frac{p^2}{4}-\alpha^2-2\alpha x^2-p\alpha$$

или

$$\left(x^2+\frac{p}{2}+\alpha\right)^2-\left[2\alpha x^2-qx+\left(\alpha^2+p\alpha-r+\frac{p^2}{4}\right)\right]=0. \tag{15}$$

Подберем теперь α так, чтобы многочлен, стоящий в квадратных скобках, стал полным квадратом. Для этого он должен иметь один двукратный корень, т. е. должно иметь место равенство

$$q^2-4\cdot 2\alpha\left(\alpha^2+p\alpha-r+\frac{p^2}{4}\right)=0. \tag{16}$$

Равенство (16) является кубичным уравнением относительно неизвестного α с комплексными коэффициентами. Это уравнение имеет, как мы знаем, три комплексных корня. Пусть α_0 будет один из них; он выражается ввиду формулы Кардано при помощи радикалов через коэффициенты уравнения (16), т. е. через коэффициенты уравнения (14).

При этом выборе значения для α многочлен, стоящий в квадратных скобках в (15), имеет двукратный корень $\frac{q}{4\alpha_0}$, и поэтому уравнение (15) принимает вид

$$\left(x^2+\frac{p}{2}+\alpha_0\right)^2-2\alpha_0\left(x-\frac{q}{4\alpha_0}\right)^2=0,$$

т. е. оно распадается на два квадратных уравнения:

$$\left.\begin{aligned} x^2-\sqrt{2\alpha_0}x+\left(\frac{p}{2}+\alpha_0+\frac{q}{2\sqrt{2\alpha_0}}\right)=0,\\ x^2+\sqrt{2\alpha_0}x+\left(\frac{p}{2}+\alpha_0-\frac{q}{2\sqrt{2\alpha_0}}\right)=0.\end{aligned}\right\} \qquad (17)$$

Так как от уравнения (14) к уравнениям (17) мы пришли при помощи тождественных преобразований, то корни уравнений (17) будут служить корнями и для уравнения (14). Легко видеть вместе с тем, что корни уравнения (14) выражаются через коэффициенты при помощи радикалов. Мы не будем выписывать соответствующих формул ввиду их громоздкости и практической бесполезности, не станем также исследовать отдельно случай, когда уравнение (14) имеет действительные коэффициенты.

Замечания об уравнениях высших степеней. В то время как методами решения квадратных уравнений владели еще древние греки, открытие изложенных выше методов решения уравнений третьей и четвертой степени относится к XVI веку. После этого почти три столетия продолжались безуспешные попытки сделать следующий шаг, т. е. найти формулы, выражающие при помощи радикалов корни любого уравнения пятой степени (т. е. уравнения пятой степени с *буквенными* коэффициентами) через его коэффициенты. Эти попытки прекратились лишь после того, как Абель в двадцатых годах прошлого века доказал, что такие формулы для уравнений n-й степени при любом $n \geqslant 5$ заведомо не могут быть найдены.

Этот результат Абеля не исключал, однако, возможности того, что корни всякого конкретного многочлена с числовыми коэффициентами все же каким-либо способом выражаются через коэффициенты при помощи некоторой комбинации радикалов, т. е., как принято говорить, что всякое уравнение разрешимо в радикалах. Полностью вопрос об условиях, при которых данное уравнение разрешимо в радикалах, был исследован Галуа в тридцатых годах прошлого века. Оказалось, что для всякого n, начиная с $n=5$, можно указать неразрешимые в радикалах уравнения n-й степени даже с целочисленными коэффициентами. Таким будет, например, уравнение

$$x^5-4x-2=0.$$

Исследования Галуа оказали решающее влияние на дальнейшее развитие алгебры. Их изложение не входит, однако, в наши задачи.

§ 39. Границы корней

Мы знаем, что не существует метода для разыскания точных значений корней многочленов с числовыми коэффициентами. Тем не менее, самые различные проблемы механики, физики и всевозможных отраслей техники сводятся к вопросу о корнях многочленов, притом иногда достаточно высоких степеней. Это обстоятельство явилось поводом для весьма многочисленных исследований, имевших целью научиться делать те или иные высказывания о корнях многочлена с числовыми коэффициентами, не зная этих корней. Изучался, например, вопрос о расположении корней на комплексной плоскости (условия, при которых все корни лежат внутри единичного круга, т. е. по модулю меньше единицы, или условия для того, чтобы все корни лежали в левой полуплоскости, т. е. имели бы отрицательные действительные части, и т. д.). Для многочленов с действительными коэффициентами разрабатывались методы определения числа их действительных корней, разыскивались границы, между которыми эти корни могут находиться, и т. д. Наконец, много исследований было посвящено методам приближенного вычисления корней: в технических приложениях обычно достаточно знать лишь приближенные значения корней с некоторой заранее данной точностью и если бы, например, корни многочлена даже записывались в радикалах, эти радикалы все равно были бы заменены их приближенными значениями.

Все эти исследования составляли в свое время основное содержание высшей алгебры. Мы включаем в наш курс лишь весьма небольшую часть относящихся сюда результатов, причем, учитывая первоочередные потребности приложений, ограничиваемся случаем многочленов с действительными коэффициентами и их действительных корней, лишь иногда выходя за эти рамки. При этом мы будем систематически рассматривать многочлен $f(x)$ с действительными коэффициентами как (непрерывную) действительную функцию действительного переменного x и всюду, где это будет полезно, будем применять результаты и методы математического анализа.

Исследование действительных корней многочлена $f(x)$ с действительными коэффициентами полезно начинать с рассмотрения графика этого многочлена: *действительными корнями многочлена будут, очевидно, абсциссы точек пересечения его графика с осью x и только они.*

Рассмотрим, например, многочлен пятой степени

$$h(x) = x^5 + 2x^4 - 5x^3 + 8x^2 - 7x - 3.$$

На основании результатов § 24 о корнях этого многочлена можно утверждать следующее: так как его степень нечетна, то $h(x)$ обладает хотя бы одним действительным корнем; если же число действительных корней больше единицы, то оно равно трем или пяти, так как комплексные корни попарно сопряжены.

Рассмотрение графика многочлена $h(x)$ позволяет сказать больше о его корнях. Построим этот график (рис. 9) [1]), беря лишь целые значения x и вычисляя соответствующие значения $h(x)$ хотя бы методом Горнера:

x	$h(x)$
.	.
.	.
.	.
—4	—39
—3	144
—2	83
—1	18
0	—3
1	—4
2	39
.	.
.	.

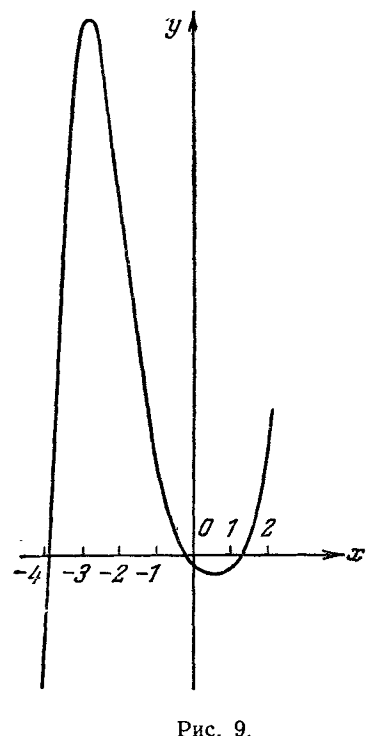


Рис. 9.

Мы видим, что многочлен $h(x)$ во всяком случае имеет три действительных корня — положительный корень α_1 и два отрицательных корня α_2 и α_3, причем

$$1 < \alpha_1 < 2, \quad -1 < \alpha_2 < 0,$$
$$-4 < \alpha_3 < -3.$$

Информация о (действительных) корнях многочлена, получающаяся от рассмотрения графика, практически обычно оказывается весьма удовлетворительной. Однако каждый раз остаются сомнения, действительно ли нами найдены все корни. Так, в рассмотренном примере мы не показали, что правее точки $x = 2$ и левее $x = -4$ уже нет корней многочлена. Больше того, так как мы брали лишь целочисленные значения x, то можно допустить, что построенный нами график не вполне точно отражает истинное поведение функции $h(x)$, не учитывает, быть может, ее более мелких колебаний и поэтому упускает некоторые корни.

Правда, можно было бы при построении графика брать не только целочисленные значения x, а значения с точностью до 0,1 или 0,01. Этим, однако, сразу чрезвычайно усложнилось бы вычисление значений $h(x)$, в то время как отмеченные выше сомнения отнюдь не были бы ликвидированы. С другой стороны, можно было бы методами

---

[1]) На рисунке масштаб по оси y взят в десять раз меньшим, чем по оси x.

математического анализа исследовать функцию $h(x)$ на максимум и минимум и таким путем сравнить наш график с истинным поведением функции; это приводит, однако, к вопросу о корнях производной $h'(x)$, т. е. к такой же задаче, как и та, которой мы занимаемся.

Отсюда вытекает потребность в более совершенных методах для разыскания границ, между которыми расположены действительные корни многочлена с действительными коэффициентами, и для определения числа этих корней. Сейчас мы будем заниматься вопросом о границах действительных корней, отнеся вопрос об их числе к следующим параграфам.

Доказательство леммы о модуле старшего члена (см. § 23) уже дает некоторую границу для модулей корней многочлена. Действительно, полагая в неравенстве (3) § 23 $k=1$, мы получаем, что при

$$|x| \geqslant 1+\frac{A}{|a_0|}, \qquad (1)$$

где a_0 — старший коэффициент, а A — максимум модулей остальных коэффициентов, модуль старшего члена многочлена больше модуля суммы всех остальных членов, а поэтому никакое значение x, удовлетворяющее неравенству (1), не может служить корнем этого многочлена.

Таким образом, *для многочлена $f(x)$ с любыми числовыми коэффициентами число $1+\frac{A}{|a_0|}$ служит верхней границей для модулей всех его корней, действительных и комплексных.* Так, для рассмотренного выше многочлена $h(x)$ этой границей, ввиду $a_0=1$, $A=8$, служит число 9.

Эта граница обычно оказывается, однако, слишком высокой, особенно если мы интересуемся лишь границами действительных корней. Сейчас будут изложены другие методы, более точные. При этом следует помнить, что если указываются границы, между которыми должны содержаться действительные корни многочлена, то этим вовсе не утверждается, что такие корни на самом деле существуют.

Покажем сначала, что *достаточно уметь находить лишь верхнюю границу положительных корней любого многочлена.* В самом деле, пусть дан многочлен $f(x)$ степени n и пусть N_0 будет верхней границей его положительных корней. Рассмотрим многочлены

$$\varphi_1(x)=x^n f\left(\frac{1}{x}\right),$$
$$\varphi_2(x)=f(-x),$$
$$\varphi_3(x)=x^n f\left(-\frac{1}{x}\right)$$

и найдем верхние границы их положительных корней; пусть это будут соответственно числа N_1, N_2, N_3. Тогда *число* $\frac{1}{N_1}$ *будет*

нижней границей положительных корней многочлена $f(x)$: *если* α есть положительный корень $f(x)$, то $\frac{1}{\alpha}$ будет положительным корнем для $\varphi_1(x)$, и из $\frac{1}{\alpha} < N_1$ следует $\alpha > \frac{1}{N_1}$. Аналогично *числа* $-N_2$ *и* $-\frac{1}{N_3}$ *служат соответственно нижней и верхней границами отрицательных корней многочлена* $f(x)$. Таким образом, все положительные корни многочлена $f(x)$ удовлетворяют неравенствам $\frac{1}{N_1} < x < N_0$, все отрицательные корни — неравенствам

$$-N_2 < x < -\frac{1}{N_3}.$$

Для определения верхней границы положительных корней можно применить следующий метод. Пусть дан многочлен

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

с действительными коэффициентами, причем $a_0 > 0$. Пусть, далее, a_k, $k \geqslant 1$, будет первым из отрицательных коэффициентов; если бы таких коэффициентов не было, то многочлен $f(x)$ вообще не мог бы иметь положительных корней. Наконец, пусть B будет наибольшая из абсолютных величин отрицательных коэффициентов. Тогда *число*

$$1 + \sqrt[k]{\frac{B}{a_0}}$$

служит верхней границей положительных корней многочлена $f(x)$.

В самом деле, полагая $x > 1$ и заменяя каждый из коэффициентов $a_1, a_2, \ldots, a_{k-1}$ числом нуль, а каждый из коэффициентов a_k, $a_{k+1}, \ldots, a_n$ — числом $-B$, мы можем лишь уменьшить значение многочлена, т. е.

$$f(x) \geqslant a_0x^n - B(x^{n-k} + x^{n-k-1} + \ldots + x + 1) = a_0x^n - B\frac{x^{n-k+1}-1}{x-1},$$

т. е., ввиду $x > 1$,

$$f(x) > a_0x^n - \frac{Bx^{n-k+1}}{x-1} = \frac{x^{n-k+1}}{x-1}[a_0x^{k-1}(x-1) - B]. \qquad (2)$$

Если

$$x > 1 + \sqrt[k]{\frac{B}{a_0}}, \qquad (3)$$

то, так как

$$a_0x^{k-1}(x-1) - B \geqslant a_0(x-1)^k - B,$$

выражение в квадратных скобках в формуле (2) окажется положительным, т. е., ввиду (2), значение $f(x)$ будет строго положительным. Таким образом, значения x, удовлетворяющие неравенству (3), не могут служить корнями для $f(x)$, что и требовалось доказать.

Для рассмотренного выше многочлена $h(x)$ этот метод дает, ввиду $k=2$ и $B=7$, в качестве верхней границы положительных корней число $1+\sqrt{7}$, что можно заменить ближайшим бо́льшим целым числом 4.

Из многочисленных других методов разыскания верхней границы положительных корней мы изложим еще лишь *метод Ньютона*. Этот метод более громоздок, чем изложенный выше, но зато дает обычно очень хороший результат.

Пусть дан многочлен $f(x)$ с действительными коэффициентами и положительным старшим коэффициентом a_0. *Если при $x=c$ многочлен $f(x)$ и все его последовательные производные $f'(x)$, $f''(x)$, ..., $f^{(n)}(x)$ принимают положительные значения, то число c служит верхней границей положительных корней.*

В самом деле, по формуле Тэйлора (см. § 23)

$$f(x)=f(c)+(x-c)f'(c)+(x-c)^2\frac{f''(c)}{2!}+\ldots+(x-c)^n\frac{f^{(n)}(c)}{n!}.$$

Мы видим, что если $x\geqslant c$, то справа будет стоять строго положительное число, т. е. такие значения x не могут служить корнями для $f(x)$.

При разыскании для данного многочлена $f(x)$ соответствующего числа c полезно поступать следующим образом. Производная $f^{(n)}(x)=n!a_0$ является положительным числом, поэтому многочлен $f^{(n-1)}(x)$ является возрастающей функцией x. Существует, следовательно, такое число c_1, что при $x\geqslant c_1$ производная $f^{(n-1)}(x)$ положительна. Отсюда следует, что при $x\geqslant c_1$ производная $f^{(n-2)}(x)$ будет возрастающей функцией x, поэтому существует такое число c_2, $c_2\geqslant c_1$, что при $x\geqslant c_2$ производная $f^{(n-2)}(x)$ также будет положительной. Продолжая далее, мы дойдем, наконец, до искомого числа c.

Применим метод Ньютона к рассматривавшемуся выше многочлену $h(x)$. Мы имеем:

$$\begin{aligned}
h(x)&=x^5+2x^4-5x^3+8x^2-7x-3,\\
h'(x)&=5x^4+8x^3-15x^2+16x-7,\\
h''(x)&=20x^3+24x^2-30x+16,\\
h'''(x)&=60x^2+48x-30,\\
h^{IV}(x)&=120x+48,\\
h^{V}(x)&=120.
\end{aligned}$$

Легко проверить (хотя бы методом Горнера), что все эти многочлены положительны при $x=2$. Таким образом, число 2 служит верхней границей положительных корней многочлена $h(x)$ — результат, много более точный, чем полученные выше другими методами.

Для разыскания нижней границы отрицательных корней многочлена $h(x)$ рассмотрим многочлен $\varphi_2(x) = -h(-x)$ [1]. Так как

$$\varphi_2(x) = x^5 - 2x^4 - 5x^3 - 8x^2 - 7x + 3,$$
$$\varphi_2'(x) = 5x^4 - 8x^3 - 15x^2 - 16x - 7,$$
$$\varphi_2''(x) = 20x^3 - 24x^2 - 30x - 16,$$
$$\varphi_2'''(x) = 60x^2 - 48x - 30,$$
$$\varphi_2^{IV}(x) = 120x - 48,$$
$$\varphi_2^{V}(x) = 120,$$

а все эти многочлены положительны, как легко проверить, при $x = 4$, то число 4 служит верхней границей положительных корней для $\varphi_2(x)$, и поэтому число -4 будет нижней границей отрицательных корней для $h(x)$.

Рассматривая, наконец, многочлены

$$\varphi_1(x) = -x^5 h\left(\frac{1}{x}\right) = 3x^5 + 7x^4 - 8x^3 + 5x^2 - 2x - 1,$$

$$\varphi_3(x) = -x^5 h\left(-\frac{1}{x}\right) = 3x^5 - 7x^4 - 8x^3 - 5x^2 - 2x + 1,$$

мы найдем для них, снова применяя метод Ньютона, в качестве верхних границ положительных корней соответственно числа 1 и 4, а поэтому нижней границей положительных корней многочлена $h(x)$ служит число $\frac{1}{1} = 1$, верхней же границей отрицательных корней — число $-\frac{1}{4}$.

Таким образом, положительные корни многочлена $h(x)$ расположены между числами 1 и 2, отрицательные корни — между числами -4 и $-\frac{1}{4}$. Этот результат очень хорошо согласуется с тем, что было найдено выше при рассмотрении графика.

§ 40. Теорема Штурма

Теперь мы перейдем к вопросу о числе действительных корней многочлена $f(x)$ с действительными коэффициентами. Мы будем при этом интересоваться как общим числом действительных корней, так и отдельно числом положительных и числом отрицательных корней и вообще числом корней, заключенных между заданными границами a и b. Существует несколько методов для разыскания точного числа корней, причем все они весьма громоздки; среди них более удобным является *метод Штурма*, который и будет сейчас изложен.

[1] Мы берем $-h(-x)$ вместо $h(-x)$ потому, что для применимости метода Ньютона старший коэффициент должен быть положительным. На корни многочлена $\varphi_2(x)$ эта перемена знака не оказывает, понятно, никакого влияния.

Введем сначала одно определение, которое будет использоваться и в следующем параграфе.

Пусть дана некоторая упорядоченная конечная система действительных чисел, отличных от нуля, например

$$1,\ 3,\ -2,\ 1,\ -4,\ -8,\ -3,\ 4,\ 1. \qquad (1)$$

Выпишем последовательно знаки этих чисел:

$$+,\ +,\ -,\ +,\ -,\ -,\ -,\ +,\ +. \qquad (2)$$

Мы видим, что в системе знаков (2) четыре раза стоят рядом противоположные знаки. Ввиду этого говорят, что в упорядоченной системе (1) имеют место четыре *перемены знаков*. Число перемен знаков можно подсчитать, понятно, для любой упорядоченной конечной системы отличных от нуля действительных чисел.

Рассмотрим теперь многочлен $f(x)$ с действительными коэффициентами, причем будем предполагать, что многочлен $f(x)$ не имеет кратных корней, так как иначе мы могли бы его разделить на наибольший общий делитель его с его производной. Конечная упорядоченная система отличных от нуля многочленов с действительными коэффициентами

$$f(x)=f_0(x),\ f_1(x),\ f_2(x),\ \ldots,\ f_s(x) \qquad (3)$$

называется *системой Штурма* для многочлена $f(x)$, если выполняются следующие требования:

1) Соседние многочлены системы (3) не имеют общих корней.

2) Последний многочлен, $f_s(x)$, не имеет действительных корней.

3) Если α — действительный корень одного из промежуточных многочленов $f_k(x)$ системы (3), $1 \leqslant k \leqslant s-1$, то $f_{k-1}(\alpha)$ и $f_{k+1}(\alpha)$ имеют разные знаки.

4) Если α — действительный корень многочлена $f(x)$, то произведение $f(x)f_1(x)$ меняет знак с минуса на плюс, когда x, возрастая, проходит через точку α.

Вопрос о том, всякий ли многочлен обладает системой Штурма, будет рассмотрен ниже; сейчас же, предполагая, что $f(x)$ такой системой обладает, покажем, как она может быть использована для нахождения числа действительных корней.

Если действительное число c не является корнем данного многочлена $f(x)$, а (3) — система Штурма для этого многочлена, то возьмем систему действительных чисел

$$f(c),\ f_1(c),\ f_2(c),\ \ldots,\ f_s(c),$$

вычеркнем из нее все числа, равные нулю, и обозначим через $W(c)$ число перемен знаков в оставшейся системе; будем называть $W(c)$

числом перемен знаков в системе Штурма (3) *многочлена* $f(x)$ *при* $x=c$[1]).

Справедлива следующая

Теорема Штурма. *Если действительные числа* a *и* b, $a<b$, *не являются корнями многочлена* $f(x)$, *не имеющего кратных корней, то* $W(a)\geqslant W(b)$ *и разность* $W(a)-W(b)$ *равна числу действительных корней многочлена* $f(x)$, *заключенных между* a *и* b.

Таким образом, для определения числа действительных корней многочлена $f(x)$, заключенных между a и b (напоминаем, что $f(x)$ по условию не имеет кратных корней), нужно лишь установить, насколько уменьшается число перемен знаков в системе Штурма этого многочлена при переходе от a к b.

Для доказательства теоремы рассмотрим, как меняется число $W(x)$ при возрастании x. Пока x, возрастая, не встретит корня ни одного из многочленов системы Штурма (3), знаки многочленов этой системы не будут меняться, и поэтому число $W(x)$ останется без изменения. Ввиду этого, а также ввиду условия 2) из определения системы Штурма, нам остается рассмотреть два случая: переход x через корень одного из промежуточных многочленов $f_k(x)$, $1\leqslant k\leqslant s-1$, и переход x через корень самого многочлена $f(x)$.

Пусть α будет корнем многочлена $f_k(x)$, $1\leqslant k\leqslant s-1$. Тогда, по условию 1), $f_{k-1}(\alpha)$ и $f_{k+1}(\alpha)$ отличны от нуля. Можно найти, следовательно, такое положительное число ε, быть может и очень малое, что в отрезке $(\alpha-\varepsilon, \alpha+\varepsilon)$ многочлены $f_{k-1}(x)$ и $f_{k+1}(x)$ не имеют корней и поэтому сохраняют постоянные знаки, причем, по условию 3), эти знаки различны. Отсюда следует, что каждая из систем чисел

$$f_{k-1}(\alpha-\varepsilon),\ f_k(\alpha-\varepsilon),\ f_{k+1}(\alpha-\varepsilon) \tag{4}$$

и

$$f_{k-1}(\alpha+\varepsilon),\ f_k(\alpha+\varepsilon),\ f_{k+1}(\alpha+\varepsilon) \tag{5}$$

обладает ровно одной переменой знаков независимо от того, каковы знаки чисел $f_k(\alpha-\varepsilon)$ и $f_k(\alpha+\varepsilon)$. Так, например, если многочлен $f_{k-1}(x)$ на рассматриваемом отрезке отрицателен, а $f_{k+1}(x)$ положителен и если $f_k(\alpha-\varepsilon)>0$, $f_k(\alpha+\varepsilon)<0$, то системам (4) и (5) соответствуют системы знаков

$$-,\ +,\ +;\ -,\ -,\ +.$$

Таким образом, при переходе x через корень одного из промежуточных многочленов системы Штурма перемены знаков в этой системе

[1]) Само собой разумеется, что перемены знаков в системе Штурма многочлена $f(x)$ не имеют ничего общего с переменой знака самого многочлена $f(x)$, происходящей от прохождения x через корень этого многочлена.

могут лишь перемещаться, но не возникают вновь и не исчезают, а поэтому число $W(x)$ при таком переходе не меняется.

Пусть, с другой стороны, α будет корнем самого данного многочлена $f(x)$. По условию 1) α не будет корнем для $f_1(x)$. Существует, следовательно, такое положительное число ε, что отрезок $(\alpha-\varepsilon, \alpha+\varepsilon)$ не содержит корней многочлена $f_1(x)$, а поэтому $f_1(x)$ сохраняет на этом отрезке постоянный знак. Если этот знак положителен, то ввиду условия 4) сам многочлен $f(x)$ при переходе x через α меняет знак с минуса на плюс, т. е. $f(\alpha-\varepsilon)<0, f(\alpha+\varepsilon)>0$. Системам чисел

$$f(\alpha-\varepsilon),\ f_1(\alpha-\varepsilon) \text{ и } f(\alpha+\varepsilon),\ f_1(\alpha+\varepsilon) \tag{6}$$

соответствуют, следовательно, системы знаков

$$-,\ + \text{ и } +,\ +,$$

т. е. в системе Штурма теряется одна перемена. Если же знак $f_1(x)$ на отрезке $(\alpha-\varepsilon, \alpha+\varepsilon)$ отрицателен, то снова, ввиду условия 4), многочлен $f(x)$ меняет знак с плюса на минус при переходе x через α, т. е. $f(\alpha-\varepsilon)>0$, $f(\alpha+\varepsilon)<0$; системам чисел (6) соответствуют теперь системы знаков

$$+,\ - \text{ и } -,\ -,$$

т. е. в системе Штурма снова теряется одна перемена.

Таким образом, *число $W(x)$ меняется (при возрастании x) лишь при переходе x через корень многочлена $f(x)$, причем в этом случае оно уменьшается ровно на единицу.*

Этим доказана, очевидно, теорема Штурма. Для того чтобы воспользоваться ею для разыскания общего числа действительных корней многочлена $f(x)$, достаточно в качестве a взять нижний предел отрицательных корней, в качестве b — верхний предел положительных корней. Проще, однако, поступить следующим образом. Ввиду леммы, доказанной в § 23, существует такое положительное число N, быть может и очень большое, что при $|x|>N$ знаки всех многочленов системы Штурма будут совпадать со знаками их старших членов. Иными словами, существует столь большое положительное значение неизвестного x, что знаки соответствующих ему значений всех многочленов системы Штурма совпадают со знаками их старших коэффициентов; это значение x, вычислять которое нет необходимости, условно обозначается символом ∞. Существует, с другой стороны, столь большое по абсолютной величине отрицательное значение x, что знаки соответствующих ему значений многочленов системы Штурма совпадают со знаками их старших коэффициентов для многочленов четной степени и противоположны знакам старших коэффициентов для многочленов нечетной степени; это значение x условимся обозначать через $-\infty$. В отрезке $(-\infty, \infty)$ содержатся, очевидно, все действительные корни всех многочленов

системы Штурма и, в частности, все действительные корни многочлена $f(x)$. Применяя к этому отрезку теорему Штурма, мы найдем число этих корней, применение же теоремы Штурма к отрезкам $(-\infty, 0)$ и $(0, \infty)$ дает соответственно число отрицательных и число положительных корней многочлена $f(x)$.

Нам остается показать, что *всякий многочлен $f(x)$ с действительными коэффициентами, не имеющий кратных корней, обладает системой Штурма*. Из различных методов, используемых для построения такой системы, мы изложим один, наиболее употребительный. Положим $f_1(x)=f'(x)$, чем обеспечивается выполнение условия 4) из определения системы Штурма. Действительно, если α — действительный корень многочлена $f(x)$, то $f'(\alpha)\neq 0$. Если $f'(\alpha)>0$, то $f'(x)>0$ в окрестности точки α, а поэтому $f(x)$ меняет знак с минуса на плюс при переходе x через α; это же верно тогда и для произведения $f(x)f_1(x)$. Аналогичные рассуждения проходят и в случае $f'(\alpha)<0$. Делим затем $f(x)$ на $f_1(x)$ и остаток от этого деления, взятый с обратным знаком, принимаем за $f_2(x)$:

$$f(x)=f_1(x)q_1(x)-f_2(x).$$

Вообще, если многочлены $f_{k-1}(x)$ и $f_k(x)$ уже найдены, то $f_{k+1}(x)$ будет остатком от деления $f_{k-1}(x)$ на $f_k(x)$, взятым с обратным знаком:

$$f_{k-1}(x)=f_k(x)q_k(x)-f_{k+1}(x). \tag{7}$$

Изложенный здесь метод отличается от алгоритма Евклида, примененного к многочленам $f(x)$ и $f'(x)$, лишь тем, что у остатка каждый раз меняется знак на обратный и следующее деление производится уже на этот остаток с обратным знаком. Так как при разыскании наибольшего общего делителя такая перемена знаков не существенна, то наш процесс остановится на некотором $f_s(x)$, являющемся наибольшим общим делителем многочленов $f(x)$ и $f'(x)$, причем из отсутствия у $f(x)$ кратных корней, т. е. из его взаимной простоты с $f'(x)$, будет следовать, что на самом деле $f_s(x)$ является некоторым отличным от нуля действительным числом.

Отсюда вытекает, что построенная нами система многочленов

$$f(x)=f_0(x),\quad f'(x)=f_1(x),\quad f_2(x),\ \ldots,\ f_s(x)$$

удовлетворяет и условию 2) из определения системы Штурма. Для доказательства выполнения условия 1) предположим, что соседние многочлены $f_k(x)$ и $f_{k+1}(x)$ обладают общим корнем α. Тогда, по (7), α будет корнем и для многочлена $f_{k-1}(x)$. Переходя к равенству

$$f_{k-2}(x)=f_{k-1}(x)q_{k-1}(x)-f_k(x),$$

мы получим, что α служит корнем и для $f_{k-2}(x)$. Продолжая далее, мы получим, что α служит общим корнем для $f(x)$ и $f'(x)$, что противоречит, однако, нашим предположениям. Наконец, выполнение

условия 3) вытекает непосредственно из равенства (7): если $f_k(\alpha)=0$, то $f_{k-1}(\alpha)=-f_{k+1}(\alpha)$.

Применим метод Штурма к рассматривавшемуся в предыдущем параграфе многочлену

$$h(x)=x^5+2x^4-5x^3+8x^2-7x-3.$$

Мы не будем при этом предварительно проверять, что $h(x)$ не имеет кратных корней, так как метод построения системы Штурма, изложенный выше, одновременно служит для проверки взаимной простоты многочлена и его производной.

Найдем систему Штурма для $h(x)$, применяя указанный метод. При этом в процессе деления мы будем, в отличие от алгоритма Евклида, умножать и сокращать лишь на произвольные положительные числа, так как знаки остатков играют в методе Штурма основную роль. Мы получим такую систему:

$$\begin{aligned}
h\ (x)&=x^5+2x^4-5x^3+8x^2-7x-3,\\
h_1(x)&=5x^4+8x^3-15x^2+16x-7,\\
h_2(x)&=66x^3-150x^2+172x+61,\\
h_3(x)&=-464x^2+1135x+723,\\
h_4(x)&=-32\,599\,457x-8\,486\,093,\\
h_5(x)&=-1.
\end{aligned}$$

Определим знаки многочленов этой системы при $x=-\infty$ и $x=\infty$, для чего, как было указано, следует смотреть лишь на знаки старших коэффициентов и на степени этих многочленов. Мы получим такую таблицу:

	$h(x)$	$h_1(x)$	$h_2(x)$	$h_3(x)$	$h_4(x)$	$h_5(x)$	Число перемен знаков
$-\infty$	$-$	$+$	$-$	$-$	$+$	$-$	4
∞	$+$	$+$	$+$	$-$	$-$	$-$	1

Таким образом, при переходе x от $-\infty$ к ∞ система Штурма теряет три перемены знаков, а поэтому многочлен $h(x)$ имеет ровно три действительных корня. Отсюда видно, что при построении в предыдущем параграфе графика этого многочлена мы не упустили ни одного из корней.

Применим метод Штурма к другому многочлену, более простому. Пусть дан многочлен

$$f(x)=x^3+3x^2-1.$$

Найдем число его действительных корней, а также целые границы, между которыми каждый из этих корней расположен, причем не будем строить заранее графика этого многочлена.

Система Штурма для многочлена $f(x)$ будет

$$\begin{aligned}
f(x)&=x^3+3x^2-1\\
f_1(x)&=3x^2+6x,\\
f_2(x)&=2x+1,\\
f_3(x)&=1.
\end{aligned}$$

Найдем число перемен знаков в этой системе при $x=-\infty$ и $x=\infty$

	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	Число перемен знаков
$-\infty$	−	+	−	+	3
∞	+	+	+	+	0

Многочлен $f(x)$ обладает, следовательно, тремя действительными корнями. Для более точного определения положения этих корней продолжим предыдущую таблицу:

	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	Число перемен знаков
$x=-3$	−	+	−	+	3
$x=-2$	+	0	−	+	2
$x=-1$	+	−	−	+	2
$x=0$	−	0	+	+	1
$x=1$	+	+	+	+	0

Таким образом, система Штурма многочлена $f(x)$ теряет по одной перемене знаков при переходе x от -3 к -2, от -1 к 0 и от 0 к 1. Корни α_1, α_2 и α_3 этого многочлена удовлетворяют, следовательно, неравенствам:

$$-3<\alpha_1<-2, \quad -1<\alpha_2<0, \quad 0<\alpha_3<1.$$

§ 41. Другие теоремы о числе действительных корней

Теорема Штурма полностью решает вопрос о числе действительных корней многочлена. Ее существенным недостатком является, однако, громоздкость вычислений, выполняемых при построении системы Штурма, как читатель мог убедиться, проделав все вычисления, относящиеся к первому из рассмотренных выше примеров. Ввиду этого сейчас будут доказаны две теоремы, не дающие точного числа действительных корней, а лишь ограничивающие это число *сверху*. Эти теоремы, применяемые после того, как при помощи графика число действительных корней уже ограничено

снизу, позволяют иногда найти точное число действительных корней, не прибегая к методу Штурма.

Пусть дан многочлен $f(x)$ n-й степени с действительными коэффициентами, причем допускаем, что он может обладать кратными корнями. Рассмотрим систему его последовательных производных

$$f(x)=f^{(0)}(x),\ f'(x),\ f''(x),\ \ldots,\ f^{(n-1)}(x),\ f^{(n)}(x), \tag{1}$$

из которых последняя равна старшему коэффициенту a_0 многочлена $f(x)$, умноженному на $n!$, и поэтому все время сохраняет постоянный знак. Если действительное число c не служит корнем ни одного из многочленов системы (1), то обозначим через $S(c)$ число перемен знаков в упорядоченной системе чисел

$$f(c),\ f'(c),\ f''(c),\ \ldots,\ f^{(n-1)}(c),\ f^{(n)}(c).$$

Таким образом можно рассматривать целочисленную функцию $S(x)$, определенную для тех значений x, которые не обращают в нуль ни одного из многочленов системы (1).

Посмотрим, как меняется число $S(x)$ при возрастании x. Пока x не пройдет через корень ни одного из многочленов (1), число $S(x)$ не может измениться. Ввиду этого мы должны рассмотреть два случая: переход x через корень многочлена $f(x)$ и переход x через корень одной из производных $f^{(k)}(x)$, $1 \leqslant k \leqslant n-1$.

Пусть α будет l-кратный корень многочлена $f(x)$, $l \geqslant 1$, т. е.

$$f(\alpha)=f'(\alpha)=\ldots=f^{(l-1)}(\alpha)=0,\quad f^{(l)}(\alpha)\neq 0.$$

Пусть положительное число ε столь мало, что отрезок $(\alpha-\varepsilon, \alpha+\varepsilon)$ не содержит корней многочленов $f(x)$, $f'(x)$, $\ldots$, $f^{(l-1)}(x)$, отличных от α, а также не содержит ни одного корня многочлена $f^{(l)}(x)$. Докажем, что в системе чисел

$$f(\alpha-\varepsilon),\ f'(\alpha-\varepsilon),\ \ldots,\ f^{(l-1)}(\alpha-\varepsilon),\ f^{(l)}(\alpha-\varepsilon)$$

всякие два соседних числа имеют противоположные знаки, тогда как все числа

$$f(\alpha+\varepsilon),\ f'(\alpha+\varepsilon),\ \ldots,\ f^{(l-1)}(\alpha+\varepsilon),\ f^{(l)}(\alpha+\varepsilon)$$

имеют один и тот же знак. Так как каждый из многочленов системы (1) является производной от предыдущего многочлена, то нам нужно лишь доказать, что если x проходит через корень α многочлена $f(x)$, то, независимо от кратности этого корня, до перехода $f(x)$ и $f'(x)$ имели разные знаки, а после перехода их знаки совпадают. Если $f(\alpha-\varepsilon)>0$, то $f(x)$ убывает на отрезке $(\alpha-\varepsilon, \alpha)$, а потому $f'(\alpha-\varepsilon)<0$; если же $f(\alpha-\varepsilon)<0$, то $f(x)$ возрастает, и потому $f'(\alpha-\varepsilon)>0$. В обоих случаях, следовательно, знаки различны. С другой стороны, если $f(\alpha+\varepsilon)>0$, то $f(x)$ возрастает на отрезке $(\alpha, \alpha+\varepsilon)$, а потому $f'(\alpha+\varepsilon)>0$; аналогично из $f(\alpha+\varepsilon)<0$

следует $f'(\alpha+\varepsilon)<0$. Таким образом, после перехода через корень α знаки $f(x)$ и $f'(x)$ должны совпадать.

Из доказанного следует, что при переходе x через l-кратный корень многочлена $f(x)$ система

$$f(x),\ f'(x),\ \ldots,\ f^{(l-1)}(x),\ f^{(l)}(x)$$

теряет l перемен знаков.

Пусть α будет теперь корнем производных

$$f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\quad 1\leqslant k\leqslant n-1,\ l\geqslant 1,$$

но не служит корнем ни для $f^{(k-1)}(x)$, ни для $f^{(k+l)}(x)$. По доказанному выше, переход x через α влечет за собой потерю в системе

$$f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\ f^{(k+l)}(x)$$

l перемен знаков. Правда, этот переход создает, возможно, новую перемену знаков между $f^{(k-1)}(x)$ и $f^{(k)}(x)$, однако, ввиду $l\geqslant 1$, при переходе x через α число перемен знаков в системе

$$f^{(k-1)}(x),\ f^{(k)}(x),\ f^{(k+1)}(x),\ \ldots,\ f^{(k+l-1)}(x),\ f^{(k+l)}(x)$$

или не меняется, или же уменьшается. Оно может уменьшиться при этом лишь на четное число, так как многочлены $f^{(k-1)}(x)$ и $f^{(k+l)}(x)$ не меняют своих знаков при переходе x через значение α.

Из полученных результатов вытекает, что *если числа a и b, $a<b$, не являются корнями ни для одного из многочленов системы* (1), *то число действительных корней многочлена $f(x)$, заключенных между a и b и подсчитываемых каждый столько раз, какова его кратность, равно разности $S(a)-S(b)$ или меньше этой разности на четное число.*

Для того чтобы ослабить ограничения, наложенные на числа a и b, введем следующие обозначения. Пусть действительное число c не является корнем многочлена $f(x)$, хотя, быть может, служит корнем для некоторых других многочленов системы (1). Обозначим через $S_+(c)$ число перемен знаков в системе чисел

$$f(c),\ f'(c),\ f''(c),\ \ldots,\ f^{(n-1)}(c),\ f^{(n)}(c), \tag{2}$$

подсчитываемое следующим образом: если

$$f^{(k)}(c)=f^{(k+1)}(c)=\ldots=f^{(k+l-1)}(c)=0, \tag{3}$$

но

$$f^{(k-1)}(c)\neq 0,\ f^{(k+l)}(c)\neq 0, \tag{4}$$

то считаем $f^{(k)}(c),\ f^{(k+1)}(c),\ \ldots,\ f^{(k+l-1)}(c)$ имеющими такой же знак, как у $f^{(k+l)}(c)$; это равносильно, очевидно, тому, что при подсчете числа перемен знаков в системе (2) нули предполагаются вычеркнутыми. С другой стороны, через $S_-(c)$ обозначим число перемен знаков в системе (2), подсчитываемое следующим образом: если имеют место условия (3) и (4), то считаем, что $f^{(k+i)}(c)$,

$0 \leqslant i \leqslant l-1$, имеет такой же знак, как и $f^{(k+l)}(c)$, если разность $l-i$ четная, и противоположный знак, если эта разность нечетная.

Если мы хотим теперь определить число действительных корней многочлена $f(x)$, заключенных между a и b, $a<b$, причем a и b не являются корнями $f(x)$, но служат, быть может, корнями для других многочленов системы (1), то поступаем следующим образом. Пусть ε столь мало, что отрезок $(a, a+2\varepsilon)$ не содержит корней многочлена $f(x)$, а также отличных от a корней всех остальных многочленов системы (1); с другой стороны, пусть η столь мало, что отрезок $(b-2\eta, b)$ также не содержит корней $f(x)$ и отличных от b корней остальных многочленов системы (1). Тогда интересующее нас число действительных корней многочлена $f(x)$ будет равно числу действительных корней этого многочлена, заключенных между $a+\varepsilon$ и $b-\eta$, т. е., по доказанному выше, равно разности $S(a+\varepsilon)-$ $-S(b-\eta)$ или меньше этой разности на четное число. Легко видеть, однако, что

$$S(a+\varepsilon)=S_{+}(a), \quad S(b-\eta)=S_{-}(b).$$

Этим доказана следующая

Теорема Бюдана—Фурье. *Если действительные числа a и b, $a<b$, не являются корнями многочлена $f(x)$ с действительными коэффициентами, то число действительных корней этого многочлена, заключенных между a и b и подсчитываемых каждый столько раз, какова его кратность, равно разности $S_{+}(a)-S_{-}(b)$ или меньше этой разности на четное число.*

Обозначим символом ∞ столь большое положительное значение неизвестного x, что знаки соответствующих ему значений всех многочленов системы (1) совпадают со знаками их старших коэффициентов. Так как этими коэффициентами будут последовательно числа a_0, na_0, $n(n-1)a_0$, ..., $n!a_0$, знаки которых совпадают, то $S(\infty)=S_{-}(\infty)=0$. С другой стороны, так как

$$f(0)=a_n, \ f'(0)=a_{n-1}, \ f''(0)=a_{n-2}2!,$$
$$f'''(0)=a_{n-3}3!, \ \ldots, \ f^{(n)}(0)=a_0 \cdot n!,$$

где $a_0, a_1, \ldots, a_n$ — коэффициенты многочлена $f(x)$, то $S_{+}(0)$ совпадает с числом перемен знаков в системе коэффициентов многочлена $f(x)$, причем коэффициенты, равные нулю, не учитываются. Таким образом, применяя теорему Бюдана—Фурье к отрезку $(0, \infty)$, мы приходим к следующей теореме:

Теорема Декарта. *Число положительных корней многочлена $f(x)$, засчитываемых каждый столько раз, какова его кратность, равно числу перемен знаков в системе коэффициентов этого многочлена (причем равные нулю коэффициенты не учитываются) или меньше этого числа на четное число.*

Для определения числа отрицательных корней многочлена $f(x)$ достаточно, очевидно, применить теорему Декарта к многочлену

$f(-x)$. При этом, если ни один из коэффициентов многочлена $f(x)$ не равен нулю, то, очевидно, переменам знаков в системе коэффициентов многочлена $f(-x)$ соответствуют сохранения знаков в системе коэффициентов многочлена $f(x)$ и обратно. Таким образом, *если многочлен $f(x)$ не имеет равных нулю коэффициентов, то число его отрицательных корней (считаемых с их кратностями) равно числу сохранений знаков в системе коэффициентов или меньше его на четное число.*

Укажем еще одно доказательство теоремы Декарта, не опирающееся на теорему Бюдана — Фурье. Докажем сначала следующую лемму:

Если $c > 0$, то число перемен знаков в системе коэффициентов многочлена $f(x)$ меньше числа перемен знаков в системе коэффициентов произведения $(x-c)f(x)$ на нечетное число.

Действительно, собирая в скобки стоящие рядом члены одного знака, запишем следующим образом многочлен $f(x)$, старший коэффициент a_0 которого считаем положительным:

$$f(x)=(a_0x^n+\ldots+b_1x^{k_1+1})-(a_1x^{k_1}+\ldots+b_2x^{k_2+1})+\ldots$$
$$\ldots+(-1)^s(a_sx^{k_s}+\ldots+b_{s+1}x^t). \quad (5)$$

Здесь $a_0>0$, $a_1>0$, ..., $a_s>0$, в то время как $b_1, b_2, \ldots, b_s$ положительны или равны нулю; однако b_{s+1} считаем строго положительным, т. е. x^t, где $t \geq 0$, является наименьшей степенью неизвестного x, входящей в многочлен $f(x)$ с отличным от нуля коэффициентом. Скобка

$$(a_0x^n+\ldots+b_1x^{k_1+1})$$

случайно может состоять при этом лишь из одного слагаемого, а именно тогда, когда $k_1+1=n$. Аналогичное замечание применимо и к другим скобкам формулы (5).

Запишем теперь многочлен, равный произведению $(x-c)f(x)$, причем будем выделять лишь члены, содержащие x в степенях $n+1$, k_1+1, ..., k_s+1 и t. Мы получим:

$$(x-c)f(x)=(a_0x^{n+1}+\ldots)-(a'_1x^{k_1+1}+\ldots)+\ldots$$
$$\ldots+(-1)^s(a'_sx^{k_s+1}+\ldots-cb_{s+1}x^t), \quad (6)$$

где $a'_i=a_i+cb_i$, $i=1, 2, \ldots, s$, и поэтому, так как $c>0$, все a'_i строго положительны. Таким образом, в системе коэффициентов многочлена $f(x)$ между членами a_0x^n и $-a_1x^{k_1}$ (а также между членами $-a_1x^{k_1}$ и $a_2x^{k_2}$ и т. д.) была одна перемена знаков, а у многочлена $(x-c)f(x)$ между соответствующими членами a_0x^{n+1} и $-a'_1x^{k_1+1}$ (соответственно между членами $-a'_1x^{k_1+1}$ и $a'_2x^{k_2+1}$ и т. д.) будет или одна перемена знаков, или больше, но тогда непременно на четное число. Точные места этих перемен знаков нас не будут

при этом интересовать; может случиться, например, что коэффициент при x^{k_1+2} в (6) отрицателен, как и коэффициент $-a_1'$, а поэтому между этими двумя соседними коэффициентами нет перемены знаков, т. е. в первой скобке перемены знаков расположены где-то раньше. Заметим теперь, что последняя скобка в (5) не содержала никаких перемен знаков, в то время как последняя скобка в (6) их содержит, притом нечетное число: достаточно учесть, что последние отличные от нуля коэффициенты многочленов $f(x)$ и $(x-c)f(x)$, т. е. $(-1)^s b_{s+1}$ и $(-1)^{s+1} b_{s+1} c$, имеют разные знаки. Таким образом, при переходе от $f(x)$ к $(x-c)f(x)$ общее число перемен знаков в системе коэффициентов непременно увеличивается, притом на нечетное число (сумма нескольких слагаемых, одно из которых нечетно, а остальные четны, будет, понятно, нечетной!). Лемма доказана.

Для доказательства теоремы Декарта обозначим через $\alpha_1, \alpha_2, \ldots, \alpha_k$ все положительные корни многочлена $f(x)$. Таким образом,

$$f(x) = (x-\alpha_1)(x-\alpha_2) \ldots (x-\alpha_k)\varphi(x),$$

где $\varphi(x)$ — многочлен с действительными коэффициентами, уже не имеющий положительных действительных корней. Отсюда следует, что первый и последний отличный от нуля коэффициенты многочлена $\varphi(x)$ одного знака, т. е. система коэффициентов этого многочлена содержит четное число перемен знаков. Применяя теперь доказанную выше лемму последовательно к многочленам

$$\varphi(x),\ (x-\alpha_1)\varphi(x),\ (x-\alpha_1)(x-\alpha_2)\varphi(x),\ \ldots,\ f(x),$$

мы получим, что число перемен знаков в системе коэффициентов каждый раз увеличивается на нечетное число, т. е. на единицу плюс четное число, а поэтому число перемен знаков в системе коэффициентов многочлена $f(x)$ больше числа k на четное число.

Применим теоремы Декарта и Бюдана—Фурье к рассматривавшемуся выше многочлену

$$h(x) = x^5 + 2x^4 - 5x^3 + 8x^2 - 7x - 3.$$

Число перемен знаков в системе коэффициентов равно трем, и поэтому, по теореме Декарта, $h(x)$ может иметь три или один положительный корень. С другой стороны, $h(x)$ не имеет равных нулю коэффициентов, а так как в системе коэффициентов два сохранения знаков, то $h(x)$ либо имеет два отрицательных корня, либо не имеет ни одного. Сравнивая с результатами, полученными ранее при помощи графика, мы получаем, что два есть точное число отрицательных корней нашего многочлена.

Для точного определения числа положительных корней воспользуемся теоремой Бюдана—Фурье, причем применим ее к отрезку $(1, \infty)$, так как в § 39 уже было показано, что 1 служит нижней границей положительных

корней многочлена $h(x)$. Последовательные производные $h(x)$ также уже были выписаны в § 39. Найдем их знаки при $x=1$ и $x=\infty$:

	$h(x)$	$h'(x)$	$h''(x)$	$h'''(x)$	$h^{IV}(x)$	$h^{V}(x)$	Число перемен знаков
$x=1$	$-$	$+$	$+$	$+$	$+$	$+$	1
$x=\infty$	$+$	$+$	$+$	$+$	$+$	$+$	0

Отсюда следует, что система производных теряет при переходе x от 1 до ∞ одну перемену знаков, а поэтому $h(x)$ имеет ровно один положительный корень.

В связи с этим примером заметим, что вообще **при разыскании числа действительных корней многочлена следует начинать с построения графика и применения теорем Декарта и Бюдана—Фурье, лишь в крайних случаях переходя к построению системы Штурма.**

Теорема Декарта допускает некоторое уточнение в том частном случае, когда заранее известно, что все корни многочлена действительные, как это имеет место, например, для характеристического многочлена симметрической матрицы. Именно:

Если все корни многочлена $f(x)$ действительные, а свободный член отличен от нуля, то число k_1 положительных корней этого многочлена равно числу s_1 перемен знаков в системе его коэффициентов, а число k_2 отрицательных корней равно числу s_2 перемен знаков в системе коэффициентов многочлена $f(-x)$.

Действительно, при наших предположениях

$$k_1+k_2=n, \tag{7}$$

где n — степень многочлена $f(x)$, и, по теореме Декарта,

$$k_1 \leqslant s_1, \quad k_2 \leqslant s_2. \tag{8}$$

Докажем, что

$$s_1+s_2 \leqslant n. \tag{9}$$

Доказательство будем вести индукцией по n, так как при $n=1$ ввиду $a_0 \neq 0$, $a_1 \neq 0$ перемена знаков имеется лишь у одного из многочленов

$$f(x)=a_0x+a_1, \quad f(-x)=-a_0x+a_1,$$

т. е. для этого случая $s_1+s_2=1$. Пусть формула (9) уже доказана для многочленов, степень которых меньше n. Если

$$f(x)=a_0x^n+a_{n-l}x^l+\ldots+a_n,$$

где $l \leqslant n-1$, $a_{n-l} \neq 0$, то положим

$$g(x) = a_{n-l}x^l + \ldots + a_n.$$

Тогда

$$f(x) = a_0x^n + g(x), \qquad f(-x) = (-1)^n a_0 x^n + g(-x).$$

Если s_1' и s_2' будут соответственно числа перемен знаков в системах коэффициентов многочленов $g(x)$ и $g(-x)$, то, по индуктивному предположению (ясно, что $l \geqslant 1$),

$$s_1' + s_2' \leqslant l.$$

Если $l = n-1$, то перемена знаков на первом месте, т. е., для $f(x)$, между a_0 и $a_1 = a_{n-l}$ будет лишь у одного из многочленов $f(x)$, $f(-x)$, а поэтому

$$s_1 + s_2 = s_1' + s_2' + 1 \leqslant l + 1 = n.$$

Если же $l \leqslant n-2$, то возможны перемены знаков на первых местах у каждого из многочленов $f(x)$, $f(-x)$, однако и в этом случае

$$s_1 + s_2 \leqslant s_1' + s_2' + 2 \leqslant l + 2 \leqslant (n-2) + 2 = n.$$

Сопоставляя (7), (8) и (9), получаем, что

$$k_1 = s_1, \quad k_2 = s_2,$$

что и требовалось доказать.

§ 42. Приближенное вычисление корней

Изложенные в предшествующих параграфах методы позволяют произвести *отделение* действительных корней многочлена $f(x)$ с действительными коэффициентами, т. е. для каждого из корней указать границы, между которыми находится только один этот корень. Если эти границы достаточно узки, то любое число, заключенное между ними, можно считать приближенным значением искомого корня. Таким образом, после того как методом Штурма (или каким-либо другим, более экономным способом) будет установлено, что между рациональными числами a и b содержится лишь один корень многочлена $f(x)$, остается задача настолько сузить эти границы, чтобы новые границы a' и b' обладали наперед заданным числом совпадающих первых десятичных знаков; этим искомый корень будет вычислен с заданной точностью.

Существует много методов, позволяющих достаточно быстро находить приближенное значение корня с требуемой точностью. Мы укажем два из них, теоретически более простые и общие и при совместном употреблении достаточно быстро приводящие к цели. Следует заметить, что методы, которые будут сейчас изложены, применимы не только к многочленам, но и к более широким классам непрерывных функций.

Будем считать дальше, что α есть простой корень многочлена $f(x)$, так как от кратных корней мы всегда можем освободиться,

и что корень α уже отделен границами a и b, $a<\alpha<b$; отсюда следует, в частности, что $f(a)$ и $f(b)$ имеют разные знаки.

Метод линейной интерполяции (называемый также методом ложного положения). В качестве приближенного значения корня α можно было бы принять, например, полусумму границ a и b, $\frac{a+b}{2}$, т. е. середину отрезка, имеющего концами a и b. Более естественно, однако, предположить, что корень лежит ближе к той из границ a, b, которой соответствует меньшее по абсолютной величине значение многочлена. Метод линейной интерполяции состоит в том, что в качестве приближенного значения корня α берется число c, делящее отрезок (a, b) на части, пропорциональные абсолютным величинам чисел $f(a)$ и $f(b)$, т. е.

$$\frac{c-a}{b-c}=-\frac{f(a)}{f(b)};$$

знак минус в правой части поставлен ввиду того, что $f(a)$ и $f(b)$ имеют разные знаки. Отсюда

$$c=\frac{bf(a)-af(b)}{f(a)-f(b)}. \tag{1}$$

Геометрически, как показывает рис. 10, метод линейной интерполяции заключается в том, что на отрезке (a, b) кривая $y=f(x)$ заменяется ее хордой, соединяющей точки $A(a, f(a))$ и $B(b, f(b))$, и в качестве приближенного значения корня α принимается абсцисса точки пересечения этой хорды с осью x.

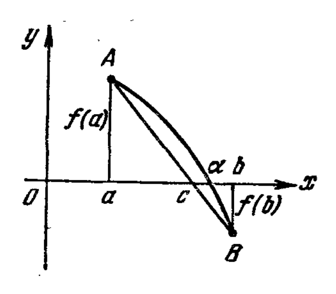


Рис. 10.

Метод Ньютона. Так как α — простой корень многочлена $f(x)$, то $f'(\alpha)\neq 0$. Примем, что также и $f''(\alpha)\neq 0$, так как иначе вопрос сводится к вычислению корня многочлена $f''(x)$, имеющего меньшую степень, чем $f(x)$. Примем, далее, что отрезок (a, b) не только не содержит корней $f(x)$, отличных от α, но и не содержит ни одного корня многочлена $f'(x)$, а также и многочлена $f''(x)$[1]. Таким образом, как следует из курса математического анализа, кривая $y=f(x)$ на отрезке (a, b) либо монотонно возрастает, либо монотонно убывает, а также либо во всех точках этого отрезка обращена выпуклостью вверх, либо во всех точках обращена выпуклостью вниз. В расположении кривой на отрезке (a, b)

[1] Сужение границ, приводящее к тому, что это условие будет удовлетворяться, достигается обычно без всяких затруднений, так как методы, изложенные ранее, позволяют установить число корней многочленов $f'(x)$ и $f''(x)$ в любом отрезке.

могут встретиться, следовательно, четыре случая, представленных на черт. 11—14.

Обозначим через a_0 тот из пределов a и b, в котором знак $f(x)$ совпадает со знаком $f''(x)$. Так как $f(a)$ и $f(b)$ имеют разные знаки, а $f''(x)$ сохраняет знак на всем отрезке (a, b), то такое a_0 может быть указано. В случаях, представленных на рис. 11 и 14,

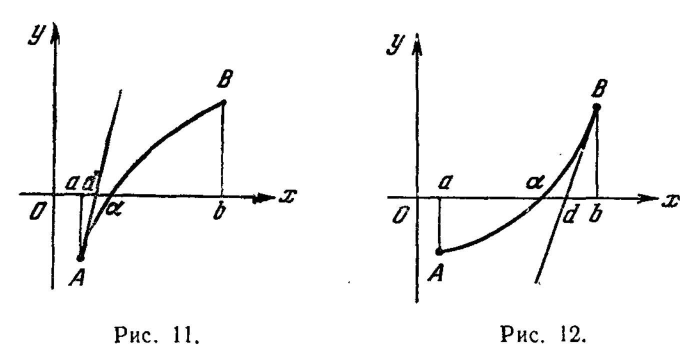


Рис. 11. Рис. 12.

будет $a_0 = a$, в двух других случаях $a_0 = b$. В точке кривой $y = f(x)$ с абсциссой a_0, т. е. в точке с координатами $(a_0, f(a_0))$, проведем касательную к этой кривой и обозначим через d абсциссу точки пересечения этой касательной с осью x. Рис. 11—14 показывают,

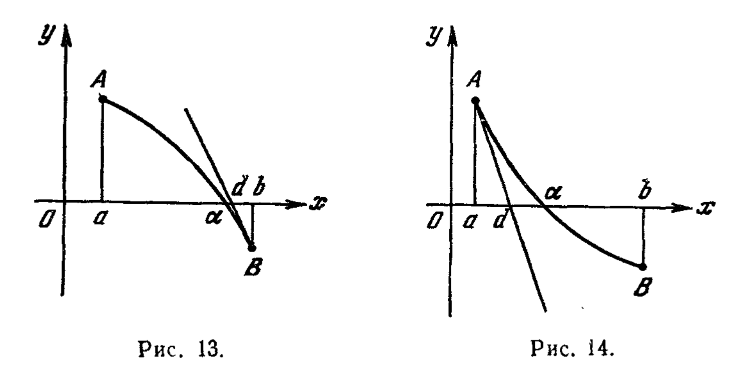


Рис. 13. Рис. 14.

что число d можно считать приближенным значением корня α. Метод Ньютона состоит, следовательно, в замене кривой $y = f(x)$ на отрезке (a, b) ее касательной в одной из границ этого отрезка. Условие, наложенное на выбор точки a_0, очень существенно: рис. 15 показывает, что без соблюдения этого условия точка пересечения касательной с осью x может вовсе не давать приближения к искомому корню.

Выведем формулу, по которой разыскивается число d. Как известно, уравнение касательной к кривой $y=f(x)$ в точке $(a_0, f(a_0))$ может быть записано в виде

$$y-f(a_0)=f'(a_0)(x-a_0).$$

Подставляя сюда координаты $(d, 0)$ точки пересечения касательной с осью x, получим:

$$-f(a_0)=f'(a_0)(d-a_0),$$

откуда

$$d=a_0-\frac{f(a_0)}{f'(a_0)}. \qquad (2)$$

Если читатель соединит на рис. 11—14 точки A и B хордами, то обнаружит, что *методы линейной интерполяции и Ньютона*

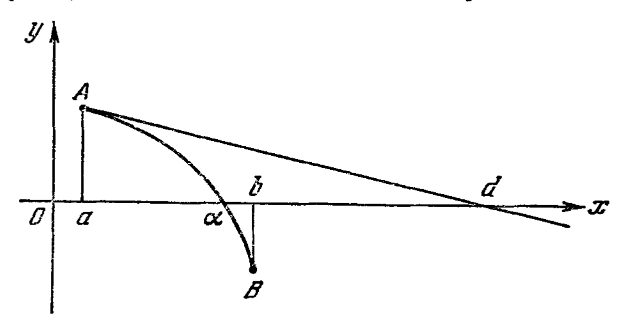


Рис. 15.

во всех случаях дают приближение к истинному значению корня α с разных сторон. Целесообразно поэтому, если отрезок (a, b) уже таков, как это требуется в методе Ньютона, комбинировать эти два метода. Мы получим этим путем много более тесные границы c и d для корня α. Если они еще не дают требуемой точности приближения, то к этим пределам следует еще раз применить указанные оба метода (см. рис. 16) и т. д., причем можно доказать, что этот процесс действительно позволяет вычислить корень α с любой точностью.

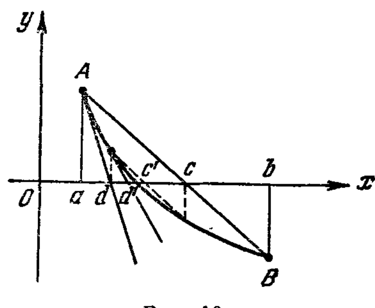


Рис. 16.

Применим эти методы к рассматривавшемуся в предшествующих параграфах многочлену

$$h(x)=x^5+2x^4-5x^3+8x^2-7x-3.$$

Мы знаем, что этот многочлен обладает простым корнем α_1, заключенным в границах $1 < \alpha_1 < 2$. Можно сказать заранее, что эти границы слишком широки для того, чтобы методы линейной интерполяции и Ньютона, примененные лишь по одному разу, могли дать хороший результат. Применим их, однако, чтобы иметь один пример, не требующий сложных вычислений.

Как мы видели в предшествующем параграфе, при $x=1$ производные $h'(x)$, $h''(x), \ldots, h^{V}(x)$ получают положительные значения. Отсюда следует, на основании результатов § 39, что значение $x=1$ служит для $h'(x)$, а также и для $h''(x)$ верхней границей положительных корней. Отрезок (1, 2) не содержит, следовательно, корней этих производных, а поэтому к нему можно применить метод Ньютона. Кроме того, $h''(x)$ всюду в этом отрезке положительна, а так как

$$h(1)=-4, \quad h(2)=39,$$

то нужно принять $a_0=2$. Учитывая, что $h'(2)=109$, мы по формуле (2) получаем:

$$d=2-\frac{39}{109}=\frac{179}{109}=1,64\ldots$$

С другой стороны, формула (1) дает:

$$c=\frac{2\cdot(-4)-1\cdot 39}{-4-39}=\frac{47}{43}=1,09\ldots$$

и, следовательно, корень α_1 заключен в границах

$$1,09 < \alpha_1 < 1,65.$$

Мы получили слишком незначительное сужение границ для того, чтобы признать этот результат удовлетворительным. Конечно, к вновь полученным границам можно было бы еще раз применить наши методы. Целесообразно, однако, с самого начала найти для α_1 достаточно тесные границы, например с точностью до 0,1 или даже 0,01, и лишь затем применять эти методы. Это сразу сделает, понятно, все вычисления весьма громоздкими, но при решении конкретных задач, требующих достаточно точного знания корней многочлена, на это приходится идти.

Вернемся к нашему многочлену $h(x)$ и его корню α_1, причем заметим, что все значения многочленов, приводимые ниже, вычисляются методом Горнера. Так как

$$h(1,3)=-0,13987, \quad h(1,31)=0,0662923851,$$

то

$$1,3 < \alpha_1 < 1,31,$$

т. е. мы нашли значение корня α_1 с точностью до 0,01. Применим теперь к этим новым границам метод линейной интерполяции:

$$c=\frac{1,31\cdot(-0,13987)-1,3\cdot 0,0662923851}{-0,13987-0,0662923851}=\frac{0,26940980063}{0,2061623851}=1,30678\ldots$$

Применим к этим же границам метод Ньютона, причем следует положить $a_0=1,31$. Так как

$$h'(1,31)=20,92822405,$$

то

$$d=1,31-\frac{0,0662923851}{20,92822405}=\frac{27,3496811204}{20,92822405}=1,30683\ldots$$

Таким образом,

$$1,30678 < \alpha_1 < 1,30684,$$

и поэтому, полагая $\alpha_1=1,30681$, мы сделаем ошибку, меньшую чем 0,00003.

Мы не показали до сих пор, что изложенные выше методы на самом деле позволяют вычислить корень с любой точностью, т. е. не доказали сходимости этих методов. Докажем это хотя бы для метода Ньютона.

Пусть, как и выше, простой корень α многочлена $f(x)$ содержится в отрезке (a, b), выбранном так, как это необходимо для применения метода Ньютона. Отсюда следует, в частности, существование таких положительных чисел A и B, что всюду на отрезке (a, b)

$$|f'(x)| > A, \quad |f''(x)| < B. \tag{3}$$

Введем обозначение

$$C = \frac{B}{2A}$$

и положим, что

$$C(b-a) < 1. \tag{4}$$

Для выполнения этого неравенства придется, возможно, заменить границы (a, b) корня α более узкими границами; это не отразится, однако, на справедливости неравенств (3). Пусть a_0 будет та из границ a, b, в которой следует применять метод Ньютона. На основании формулы (2) мы последовательно получим в качестве приближенных значений корня α числа $a_1, a_2, \ldots, a_k, \ldots$, лежащие в отрезке (a, b) и связанные между собой равенствами

$$a_k = a_{k-1} - \frac{f(a_{k-1})}{f'(a_{k-1})}, \quad k = 1, 2, \ldots \tag{5}$$

Пусть

$$\alpha = a_k + h_k, \quad k = 0, 1, 2, \ldots \tag{6}$$

Тогда

$$0 = f(\alpha) = f(a_k) + h_k f'(a_k) + \frac{h_k^2}{2} f''(a_k + \theta h_k),$$

где $0 < \theta < 1$. Так как $f'(a_k) \neq 0$ ввиду условия, наложенного на отрезок (a, b), то, учитывая (5) и (6), получим:

$$-\frac{h_k^2}{2} \frac{f''(a_k+\theta h_k)}{f'(a_k)} = h_k + \frac{f(a_k)}{f'(a_k)} = \alpha - \left(a_k - \frac{f(a_k)}{f'(a_k)}\right) = \alpha - a_{k+1} = h_{k+1}.$$

Отсюда

$$|h_{k+1}| = h_k^2 \left| \frac{f''(a_k+\theta h_k)}{2f'(a_k)} \right| < h_k^2 \frac{B}{2A} = C h_k^2, \quad k = 0, 1, 2, \ldots$$

Таким образом,

$$|h_{k+1}| < C h_k^2 < C^3 h_{k-1}^4 < C^7 h_{k-2}^8 < \ldots < C^{2^{k+1}-1} h_0^{2^{k+1}}$$

или, так как $|h_0| = |\alpha - a_0| < b - a$,

$$|h_{k+1}| < C^{-1} [C(b-a)]^{2^{k+1}}, \quad k = 0, 1, 2, \ldots \tag{7}$$

Отсюда, ввиду условия (4), следует, что *разность h_k между корнем α и его приближенным значением a_k, полученным последовательным применением метода Ньютона, стремится к нулю при возрастании k*, что и требовалось доказать.

Отметим, что формула (7) дает *оценку погрешности* для $(k+1)$-го шага, что существенно, если метод Ньютона применяется один, а не в комбинации с методом линейной интерполяции.

В курсах теории приближенных вычислений читатель может познакомиться со способами более рационального расположения вычислений в изложенных выше методах, облегчающими их применение. В этих же курсах можно найти изложение многих других методов приближенного вычисления корней. Среди них наиболее совершенным является *метод Лобачевского* (иногда ошибочно называемый методом Греффе). Этот метод позволяет находить приближенные значения всех корней сразу, в том числе и комплексных, причем не требует предварительного отделения корней; он связан, однако, с весьма громоздкими вычислениями. В основе этого метода лежит излагаемая ниже, в гл. 11, теория симметрических многочленов.

ГЛАВА ДЕСЯТАЯ

ПОЛЯ И МНОГОЧЛЕНЫ

§ 43. Числовые кольца и поля

В очень многих предшествующих разделах курса мы оказывались в следующем положении: излагая материал, мы допускали к рассмотрению или любые комплексные числа, или же только действительные числа, но затем должны были делать замечание, что полученные результаты остаются справедливыми, если ограничиться лишь действительными числами (или, соответственно, что они дословно переносятся на случай любых комплексных чисел). Как правило, во всех этих случаях можно было заметить, что изложенная теория полностью сохранилась бы и в том случае, если бы мы допустили к рассмотрению лишь рациональные числа. Настало время показать читателю истинные причины этого параллелизма с тем, чтобы излагать дальнейший материал в естественной для него общности, т. е. на общепринятом алгебраическом языке. С этой целью мы введем понятие поля, а также более широкое, но в нашем курсе играющее лишь служебную роль, понятие кольца.

Очевидно, что системы всех комплексных, всех действительных и всех рациональных чисел, равно как и система всех целых чисел, *обладают тем общим свойством, что в каждой из них не только сложение и умножение, но и вычитание можно выполнять, оставаясь в пределах самой этой системы.* Это свойство указанных числовых систем отличает их, например, от системы положительных целых или положительных действительных чисел.

Всякая система чисел, комплексных или, в частности, действительных, содержащая сумму, разность и произведение любых двух своих чисел, называется *числовым кольцом.* Таким образом, системы всех целых, рациональных, действительных и комплексных чисел являются числовыми кольцами. С другой стороны, никакая система положительных чисел не будет кольцом, так как если a и b — два различных положительных числа, то либо $a-b$, либо $b-a$ отрицательно. Не будет кольцом и никакая система отрицательных чисел хотя бы потому, что произведение двух отрицательных чисел положительно.

Числовые кольца далеко не исчерпываются рассмотренными выше четырьмя примерами. Сейчас будут указаны некоторые другие при-

меры, причем проверка утверждения, что рассматриваемая система чисел действительно является кольцом, каждый раз предоставляется читателю.

Четные числа составляют кольцо; вообще при любом натуральном n совокупность целых чисел, нацело делящихся на n, будет кольцом. Нечетные числа кольца не составляют, так как сумма двух нечетных чисел четна.

Кольцом будет совокупность рациональных чисел, знаменатели записей которых в виде несократимой дроби являются какими-либо степенями числа 2; к этой совокупности принадлежат, в частности, все целые числа, так как их несократимые записи имеют знаменателем число 1, т. е. два в нулевой степени. В этом примере вместо числа 2 можно взять, конечно, любое простое число p. Вообще, беря любое множество простых чисел, конечное или даже бесконечное, и рассматривая систему рациональных чисел, знаменатели несократимых записей которых могут делиться лишь на простые числа, принадлежащие к взятому множеству, мы также получим кольцо. С другой стороны, совокупность рациональных чисел, знаменатели несократимых записей которых не делятся на квадрат никакого простого числа, не будет кольцом, так как указанное свойство чисел не сохраняется при их умножении.

Переходим к примерам числовых колец, не лежащих целиком в кольце рациональных чисел. Совокупность чисел вида

$$a+b\sqrt{2}, \tag{1}$$

где a и b — любые рациональные числа, будет кольцом; к этому кольцу принадлежат, в частности, все рациональные числа (при $b=0$), а также само число $\sqrt{2}$ (при $a=0$, $b=1$). Мы получили бы также кольцо, если бы ограничились лишь числами вида (1) с целыми коэффициентами a, b. В этих примерах можно, конечно, вместо числа $\sqrt{2}$ взять $\sqrt{3}$ или $\sqrt{5}$ и т. д.

Система чисел вида

$$a+b\sqrt[3]{2} \tag{2}$$

с любыми рациональными (или лишь с любыми целыми) коэффициентами a, b не будет кольцом, так как произведение числа $\sqrt[3]{2}$ на самого себя нельзя, как легко проверить, записать в виде (2) [1]. Однако система чисел вида

$$a+b\sqrt[3]{2}+c\sqrt[3]{4} \tag{3}$$

с любыми рациональными коэффициентами a, b, c уже будет кольцом, и это же имеет место, если ограничиться случаем целых коэффициентов.

[1]) Действительно, пусть

$$\sqrt[3]{4}=a+b\sqrt[3]{2}, \tag{2'}$$

где числа a и b рациональны. Умножая обе части этого равенства на $\sqrt[3]{2}$,

Рассмотрим теперь все действительные числа, которые можно получить, применяя несколько раз операции сложения, умножения и вычитания к хорошо известному читателю числу π и каким-либо рациональным числам. Это будут числа, которые могут быть записаны в виде

$$a_0+a_1\pi+a_2\pi^2+\ldots+a_n\pi^n, \qquad (4)$$

где $a_0, a_1, a_2, \ldots, a_n$ — рациональные числа, $n \geqslant 0$. Заметим, что никакое число не может обладать двумя различными записями вида (4) — в противном случае, беря разность двух таких записей, мы получили бы, что число π удовлетворяет некоторому уравнению с рациональными коэффициентами; методами математического анализа доказывается, однако, что π не может удовлетворять на самом деле никакому уравнению с рациональными коэффициентами, т. е. является числом трансцендентным. Не используя, впрочем, этого результата, т. е. не предполагая, что запись числа в виде (4) однозначна, можно все же показать, что числа вида (4) составляют кольцо.

Кольцом будет также совокупность чисел, получающихся из числа π и рациональных чисел при помощи операций сложения, умножения, вычитания и деления, примененных несколько раз. Для доказательства нет необходимости искать для рассматриваемых чисел какую-либо специальную хорошую запись (хотя она и может быть найдена): если числа α и β получены из числа π и некоторых рациональных чисел указанными операциями, то это же верно, понятно, и для чисел $\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$, а также (при $\beta\neq 0$) для числа $\frac{\alpha}{\beta}$.

Наконец, взяв совокупность комплексных чисел $a+bi$ с любыми рациональными a, b, мы получим кольцо; это же будет иметь место, если мы ограничимся целыми коэффициентами a, b.

Рассмотренные примеры не могут дать полного представления о том, сколь разнообразными бывают числовые кольца. Мы не будем пока, однако, продолжать наш список примеров и перейдем к рассмотрению одного специального и очень важного типа числовых

получим:

$$2=a\sqrt[3]{2}+b\sqrt[3]{4}.$$

Подставляя сюда выражение (2′) для $\sqrt[3]{4}$, мы после очевидных преобразований придем к равенству

$$(a+b^2)\sqrt[3]{2}=2-ab. \qquad (2'')$$

Если $a+b^2\neq 0$, то

$$\sqrt[3]{2}=\frac{2-ab}{a+b^2},$$

что невозможно, так как справа стоит рациональное число. Если же $a+b^2=0$, то, ввиду (2″), и $2-ab=0$. Из этих двух равенств вытекает $b^3=-2$, что снова невозможно ввиду рациональности числа b.

колец. Мы знаем, конечно, что в системах всех рациональных, всех действительных и всех комплексных чисел можно неограниченно выполнять деление (кроме деления на нуль), в то время как деление целых чисел выводит за пределы системы этих чисел. До сих пор мы не обращали серьезного внимания на это различие, в действительности же оно очень существенно и приводит к следующему определению.

Числовое кольцо называется *числовым полем*, если оно содержит частное любых двух своих чисел (делитель предполагается, конечно, отличным от нуля). Можно говорить, следовательно, о поле рациональных чисел, поле действительных чисел, полекомплексных чисел, в то время как кольцо целых чисел полем не является.

Некоторые из рассмотренных выше примеров числовых колец в действительности являются полями. Сначала заметим, что не существует числовых полей, отличных от поля рациональных чисел и целиком в нем содержащихся (систему, состоящую из одного нуля, мы не будем считать полем). Справедливо даже следующее более общее утверждение:

Поле рациональных чисел содержится целиком во всяком числовом поле.

Пусть, в самом деле, дано некоторое числовое поле, которое мы обозначим буквой P. Если a — любое число поля P, отличное от нуля, то P содержит и частное от деления числа a на самого себя, т. е. число единицу. Складывая единицу с самой собою несколько раз, мы получим, что все натуральные числа содержатся в поле P. С другой стороны, в поле P должна содержаться разность $a-a$, т. е. число нуль, а поэтому к P принадлежит и результат вычитания любого натурального числа из нуля, т. е. любое целое отрицательное число. Наконец, в поле P лежат и частные целых чисел, т. е. вообще все рациональные числа.

В поле комплексных чисел содержится много различных полей, и поле рациональных чисел будет лишь наименьшим среди них. Так, рассмотренное выше кольцо чисел вида

$$a+b\sqrt{2} \tag{5}$$

с любыми рациональными (а не только лишь с целыми) коэффициентами a, b будет полем. В самом деле рассмотрим частное двух чисел вида (5), $a+b\sqrt{2}$ и $c+d\sqrt{2}$, причем второе число считаем отличным от нуля; отлично от нуля, следовательно, и число $c-d\sqrt{2}$, и потому

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}.$$

Мы получили снова число вида (5), причем коэффициенты остаются рациональными. В этом примере число $\sqrt{2}$ можно заменить,

понятно, квадратным корнем из любого рационального числа, из которого в самом поле рациональных чисел не извлекается квадратный корень. Так, поле составляют числа вида $a+bi$ с рациональными a, b.

§ 44. Кольцо

В различных отделах математики, а также в применениях математики к технике и естествознанию приходится весьма часто встречаться с положением, когда алгебраические операции производятся не над числами, а над объектами совсем иной природы. Большое число таких примеров можно найти в предшествующих главах книги — напомним умножение и сложение матриц, сложение векторов, операции над многочленами, операции над линейными преобразованиями. Общее определение *алгебраической операции*, которому удовлетворяют операции сложения и умножения в числовых кольцах, а также операции в указанных примерах, состоит в следующем.

Пусть дано некоторое множество M, состоящее или из чисел, или из объектов геометрической природы, вообще из некоторых вещей, которые мы будем называть *элементами* этого множества. Говорят, что *в множестве M определена алгебраическая операция*, если указан закон, по которому любой паре элементов a, b из этого множества однозначным образом ставится в соответствие некоторый третий элемент c, также принадлежащий к M. Эта операция может быть названа *сложением*, и тогда c будет называться *суммой* элементов a и b и обозначаться символом $c=a+b$; эта операция может быть названа *умножением*, т. е. c будет *произведением* элементов a и b, $c=ab$; возможно, наконец, что для операции, определенной в множестве M, будет введена новая терминология и символика.

В каждом из числовых колец определены две независимые операции — сложение и умножение. Что же касается вычитания и деления, то их нельзя считать новыми операциями, так как они являются обратными соответственно для сложения и для умножения, если мы примем следующее общее определение *обратной операции*.

Пусть в множестве M определена алгебраическая операция, например сложение. Говорят, что *для этой операции существует обратная операция* — вычитание, если для любой пары элементов a, b из M существует в M такой элемент d, притом лишь единственный, который удовлетворяет равенству $b+d=a$. Элемент d называется тогда *разностью* элементов a и b и обозначается символом $d=a-b$.

В числовых полях обратной операцией обладает, очевидно, как сложение, так и умножение (последнее, правда, ограниченно: делитель должен быть отличным от нуля). В числовых же кольцах, не являющихся полями (как, например, в кольце целых чисел), обратной операцией обладает лишь сложение.

С другой стороны, в системе всех многочленов от неизвестного x, коэффициенты которых принадлежат к фиксированному числовому полю P, также определены две операции — сложение и умножение, причем сложение обладает обратной операцией — вычитанием.

И в числовых кольцах, и в системе многочленов операции сложения и умножения обладают, как известно, следующими свойствами (a, b, c — произвольные числа из рассматриваемого числового кольца или произвольные многочлены из рассматриваемой системы):

I. Сложение коммутативно: $a+b=b+a$.

II. Сложение ассоциативно: $a+(b+c)=(a+b)+c$.

III. Умножение коммутативно: $ab=ba$.

IV. Умножение ассоциативно: $a(bc)=(ab)c$.

V. Сложение и умножение связаны законом дистрибутивности:

$$(a+b)c=ac+bc.$$

Мы уже подготовлены теперь к общему определению понятия кольца, одного из важнейших понятий алгебры.

Множество R называется *кольцом*, если в нем определены две операции — сложение и умножение, обе коммутативные и ассоциативные, а также связанные законом дистрибутивности, причем сложение обладает обратной операцией — вычитанием.

Таким образом, примерами колец являются числовые кольца и кольца многочленов от неизвестного x с коэффициентами из данного числового поля или даже из данного числового кольца. Укажем еще один пример, хорошо выясняющий широту понятия кольца.

Курс математического анализа начинается с определения функции действительного переменного x. Рассмотрим совокупность функций, определенных для всех действительных значений x и принимающих действительные значения, и следующим образом определим в этой совокупности алгебраические операции: *суммой* двух функций $f(x)$ и $g(x)$ будет функция, значение которой при любом $x=x_0$ равно сумме значений заданных функций, т. е. равно $f(x_0)+g(x_0)$, *произведением* этих функций — функция, значение которой при всяком $x=x_0$ равно произведению $f(x_0)\cdot g(x_0)$. Сумма и произведение существуют, очевидно, для любых двух функций из рассматриваемой совокупности. Справедливость свойств I—V проверяется без всяких затруднений — сложение и умножение функций сводятся к сложению и умножению их значений при всяком x, т. е. к операциям над действительными числами, для которых свойства I—V имеют место. Наконец, считая *разностью* функций $f(x)$ и $g(x)$ функцию, значение которой при любом $x=x_0$ равно разности $f(x_0)-g(x_0)$, мы придем к операции вычитания, обратной сложению. Этим доказано, что *совокупность функций, определенных для всех действительных x, после введения в нее описанным выше способом операций сложения и умножения превращается в кольцо.*

Другие примеры колец функций можно получить, сохраняя данные выше определения операций над функциями, но рассматривая функции, определенные, например, лишь для положительных значений переменного x, или функции, определенные для значений x из отрезка $[0, 1]$. Вообще кольцом будет система всех функций, имеющих некоторую данную область определения. Можно было бы получить также примеры колец, рассматривая не все функции, определенные в данной области, а лишь изучаемые в курсе математического анализа непрерывные функции. Можно было бы, с другой стороны, рассматривать комплексные функции комплексного переменного. Вообще, различных колец функций, как и различных числовых колец, чрезвычайно много.

Переходим к установлению некоторых простейших свойств колец, непосредственно вытекающих из определения кольца. Эти свойства для случая чисел вполне привычны, однако читателю, быть может, покажется иногда неожиданным, что они являются следствиями лишь условий I—V и существования однозначного вычитания.

Сначала несколько замечаний о значении условий I—V. Роль *законов коммутативности* не требует пояснений. Значение *законов ассоциативности* состоит в следующем: в определении алгебраической операции говорится о сумме или произведении лишь двух элементов. Если же мы попытаемся определить, например, произведение трех элементов a, b, c, то встретимся с таким затруднением: произведения au и vc, где $bc=u$, $ab=v$, могут, вообще говоря, не совпадать, т. е. $a(bc)\neq(ab)c$. Закон ассоциативности требует, чтобы эти произведения были равны одному и тому же элементу кольца: этот элемент естественно принять в качестве произведения abc, записываемого уже без всяких скобок. Больше того, *закон ассоциативности позволяет однозначным образом определить произведение* (*соответственно*, *сумму*) *для любого конечного числа элементов кольца*, т. е. позволяет доказать независимость произведения любых n элементов от первоначального распределения скобок.

Докажем это утверждение индукцией по числу n. Для $n=3$ оно уже доказано, поэтому полагаем $n>3$, причем считаем, что для всех чисел, меньших n, наше утверждение уже доказано. Пусть даны элементы $a_1, a_2, \ldots, a_n$ и пусть в этой системе некоторым образом распределены скобки, указывающие на порядок, в каком должно выполняться умножение. Последним шагом будет умножение произведения первых k элементов $a_1a_2\ldots a_k$ (где $1\leqslant k\leqslant n-1$) на произведение $a_{k+1}a_{k+2}\ldots a_n$. Так как эти произведения состоят из меньшего, чем n, числа множителей и поэтому, по предположению, однозначно определены, то нам остается доказать для любых k и l равенство

$$(a_1a_2\ldots a_k)(a_{k+1}a_{k+2}\ldots a_n)=(a_1a_2\ldots a_l)(a_{l+1}a_{l+2}\ldots a_n).$$

Для этого достаточно рассмотреть случай $l=k+1$. В этом случае, однако, полагая

$$a_1a_2\ldots a_k=b,\quad a_{k+2}a_{k+3}\ldots a_n=c,$$

мы получаем на основании закона ассоциативности

$$b(a_{k+1}c)=(ba_{k+1})c.$$

Этим наше утверждение доказано.

Можно говорить, в частности, о произведении n равных между собою элементов, т. е. ввести понятие о *степени* a^n элемента a с целым положительным показателем n. Легко проверить, что все обычные правила оперирования с показателями остаются справедливыми в любом кольце. Закон ассоциативности сложения приводит аналогичным образом к понятию о *кратном* na элемента a с целым положительным коэффициентом n.

Закон дистрибутивности, т. е. обычное правило раскрытия скобок, является единственным требованием в определении кольца, связывающим сложение и умножение; лишь благодаря этому закону совместное изучение двух указанных операций дает больше, чем можно было бы получить при их раздельном изучении. В формулировке закона дистрибутивности участвует сумма лишь двух слагаемых. Без всякого труда доказывается, однако, справедливость равенства

$$(a_1+a_2+\ldots+a_k)\,b=a_1b+a_2b+\ldots+a_kb$$

при любом k, а затем и общего правила умножения суммы на сумму.

Во всяком кольце выполняется закон дистрибутивности и для разности. Действительно, по определению разности элемент $a-b$ удовлетворяет равенству

$$b+(a-b)=a.$$

Умножая обе части этого равенства на c и применяя к левой части равенства закон дистрибутивности, мы получаем:

$$bc+(a-b)\,c=ac.$$

Элемент $(a-b)\,c$ является, следовательно, разностью элементов ac и bc:

$$(a-b)\,c=ac-bc.$$

Весьма важные свойства колец вытекают из существования вычитания. Если a есть произвольный элемент кольца R, то разность $a-a$ будет некоторым вполне определенным элементом кольца. Его роль аналогична роли нуля в числовых кольцах, однако по определению он может зависеть от выбора элемента a, и поэтому мы обозначим его пока через 0_a.

Докажем, что на самом деле элементы 0_a для всех a равны между собой. Действительно, если b есть произвольный другой элемент кольца R, то, прибавляя к обеим частям равенства

$$a+(b-a)=b$$

элемент 0_a и используя равенство $0_a+a=a$, мы получаем:

$$0_a+b=0_a+a+(b-a)=a+(b-a)=b.$$

Таким образом, $0_a=b-b=0_b$.

Мы доказали, *что всякое кольцо R обладает однозначно определенным элементом, сумма которого с любым элементом a этого кольца равна a.* Будем называть этот элемент *нулем* кольца R и обозначать символом 0, не считая серьезной опасность смешать его с числом нуль. Таким образом,

$$a+0=a \text{ для всех } a \text{ из } R.$$

Далее, *во всяком кольце для любого элемента a существует однозначно определенный противоположный элемент $-a$, удовлетворяющий равенству*

$$a+(-a)=0,$$

а именно, этим элементом будет разность $0-a$; однозначность вытекает из однозначности вычитания. Очевидно, что $-(-a)=a$. Разность $b-a$ двух любых элементов кольца можно записать теперь в виде

$$b-a=b+(-a).$$

Действительно,

$$[b+(-a)]+a=b+[(-a)+a]=b+0=b.$$

Для любого элемента a кольца и любого целого положительного числа n имеет место равенство

$$n(-a)=-(na).$$

Действительно, группировкой слагаемых получаем:

$$na+n(-a)=n[a+(-a)]=n\cdot 0=0.$$

Мы получили теперь возможность определить *отрицательные кратные* элемента кольца: если $n>0$ то равные между собой элементы $n(-a)$ и $-(na)$ будут обозначаться через $(-n)a$. Условимся, наконец, *нулевым кратным* $0\cdot a$ любого элемента a считать нуль рассматриваемого кольца.

Определение нуля дано нами лишь при помощи операции сложения и ей обратной, т. е. без использования умножения. В случае чисел, однако, число нуль и по отношению к умножению обладает одним характерным и притом очень важным свойством. Оказывается, что этим свойством обладает нуль любого кольца: *во всяком кольце произведение любого элемента на нуль равно нулю.* Доказательство непосредственно опирается на закон дистрибутивности: если a есть произвольный элемент кольца R, то, каков бы ни был вспомогательный элемент x этого кольца, мы получим:

$$a\cdot 0=a(x-x)=ax-ax=0.$$

Пользуясь этим свойством нуля, можно доказать, что *во всяком кольце для любых элементов a, b справедливо равенство*

$$(-a)b=-ab.$$

Действительно,

$$ab+(-a)b=[a+(-a)]b=0\cdot b=0.$$

Отсюда следует, что хорошо известное и все же несколько таинственное правило умножения отрицательных чисел — «минус на минус дает плюс» — также вытекает из определения кольца, т. е. *в любом кольце имеет место равенство*

$$(-a)(-b)=ab.$$

В самом деле,

$$(-a)(-b)=-[a(-b)]=-(-ab)=ab.$$

Читатель без труда докажет теперь, что во всяком кольце для кратных (в том числе и отрицательных) любого элемента остаются справедливыми все правила оперирования с кратными некоторого числа.

Таким образом, алгебраические операции в произвольном кольце обладают многими привычными нам свойствами операций над числами. Не следует думать, однако, что любое свойство сложения и умножения чисел сохраняется во всяком кольце. Так, умножение чисел обладает свойством, обратным рассмотренному выше: если произведение двух чисел равно нулю, то хотя бы один из множителей равен нулю. Это свойство уже не может быть распространено на любые кольца — в некоторых кольцах можно указать такие пары отличных от нуля элементов, произведение которых равно нулю, т. е. $a\neq 0$, $b\neq 0$, но $ab=0$; элементы a, b с этим свойством называются *делителями нуля.*

Примеров колец с делителями нуля нельзя найти, понятно, среди числовых колец. Не содержат делителей нуля и кольца многочленов с числовыми коэффициентами. Многие кольца функций обладают, однако, делителями нуля. Заметим, прежде всего, что нулем во всяком кольце функций будет функция, равная нулю при всех значениях переменного x. Построим теперь следующие функции $f(x)$ и $g(x)$, определенные для всех действительных значений x:

$$f(x)=0 \quad \text{при} \quad x\leqslant 0, \quad f(x)=x \quad \text{при} \quad x>0;$$
$$g(x)=x \quad \text{при} \quad x\leqslant 0, \quad g(x)=0 \quad \text{при} \quad x>0.$$

Обе эти функции отличны от нуля, так как не при всех значениях x равны нулю их значения; произведение же этих функций равно нулю.

Не все требования I — V, входящие в определение кольца, являются в одинаковой мере необходимыми. Развитие науки показывает, что в то время как свойства сложения I и II и закон дистрибутивности V имеют место во всех приложениях, включение в определение кольца свойств умножения III и IV часто оказывается излишне стеснительным, суживая возможную область применимости этого понятия. Так, множество квадратных матриц порядка n с действительными элементами, рассматриваемое с операциями сложения и умножения матриц, удовлетворяет всем требованиям, входящим в определение кольца, за исключением закона коммутативности умножения.

С некоммутативными умножениями приходится встречаться так часто и в таких важных случаях, что в настоящее время под термином «кольцо» понимают обычно *некоммутативное кольцо* (точнее, не обязательно коммутативное кольцо, в смысле возможной некоммутативности умножения), называя тот частный тип колец, в которых требование III выполняется, *коммутативными кольцами*.

В последнее время повышается интерес и к кольцам с неассоциативным умножением и общая теория колец уже строится сейчас как теория неассоциативных (т. е. не обязательно ассоциативных) колец. Простейшим примером таких колец является множество векторов трехмерного евклидова пространства относительно операций сложения и (известного из курса аналитической геометрии) векторного умножения векторов.

§ 45. Поле

Подобно тому как среди числовых колец были выделены и названы числовыми полями те кольца, в которых можно выполнять деление (кроме деления на нуль), естественно сделать это и в общем случае. Заметим сначала, что *ни в каком кольце невозможно деление на нуль* ввиду доказанного выше свойства нуля по отношению к умножению: разделить элемент a на нуль означает найти в кольце такой элемент x, что $0 \cdot x = a$, что при $a \neq 0$ невозможно, так как левая часть равна нулю.

Введем следующее определение:

Кольцо P называется *полем*, если оно состоит не только из одного нуля и если в нем деление выполнимо, притом однозначным образом, во всех случаях, кроме случая деления на нуль, т. е. если для любых элементов a и b из P, из которых b отлично от нуля, существует в P такой элемент q, притом лишь единственный, который удовлетворяет равенству $bq = a$. Элемент q называется *частным* элементов a и b и обозначается символом $q = \frac{a}{b}$ [1].

Примерами полей служат, понятно, все числовые поля. Кольцо многочленов от неизвестного x с действительными коэффициентами или вообще с коэффициентами из некоторого числового поля не является полем — существующее для многочленов деление с остатком отличается, конечно, от деления «нацело», предполагающегося в определении поля. С другой стороны, легко видеть, что *совокупность всех дробно-рациональных функций с действительными коэффициентами* (см. § 25) *будет полем*, содержащим кольцо многочленов, подобно тому как поле рациональных чисел содержит кольцо целых чисел.

Среди колец функций можно указать некоторые другие примеры полей; мы не будем, однако, на них останавливаться и перейдем к примерам совсем иного рода.

---

[1]) Единственность деления в поле, как и предполагавшаяся в определении кольца единственность вычитания, в действительности без труда могут быть доказаны при помощи других требований, входящих в определение поля или, соответственно, кольца.

Все числовые кольца и вообще кольца, которые мы до сих пор рассматривали, содержат бесконечно много элементов. Существуют, однако, кольца и даже поля, состоящие лишь из конечного числа элементов. Простейшие примеры *конечных колец* и *конечных полей*, существенно используемые в особой ветви математики — теории чисел, строятся следующим образом.

Берем любое натуральное число n, отличное от 1. Целые числа a и b называются *сравнимыми по модулю* n,

$$a \equiv b \pmod{n},$$

если эти числа дают при делении на n один и тот же остаток, т. е. если их разность нацело делится на n. Все кольцо целых чисел распадается на n непересекающихся классов,

$$C_0, \ C_1, \ \ldots, \ C_{n-1}, \qquad (1)$$

сравнимых между собой по модулю n чисел, причем класс C_k, $k=0, 1, \ldots, n-1$, состоит из чисел, дающих при делении на n в остатке k. Оказывается, что можно вполне естественным способом определить сложение и умножение этих классов.

Возьмем с этой целью любые (притом не обязательно различные) классы C_k и C_l из системы (1). Складывая любое число из класса C_k с любым числом из класса C_l, мы будем получать числа, лежащие в одном вполне определенном классе, а именно в классе C_{k+l}, если $k+l<n$, или в классе C_{k+l-n}, если $k+l \geqslant n$. Это приводит к такому определению *сложения классов*:

$$\begin{aligned} C_k + C_l &= C_{k+l} \quad \text{при} \quad k+l<n, \\ C_k + C_l &= C_{k+l-n} \quad \text{при} \quad k+l \geqslant n. \end{aligned} \qquad (2)$$

С другой стороны, умножая любое число класса C_k на любое число класса C_l, мы будем получать числа, снова лежащие во вполне определенном классе, а именно в классе C_r, где r — остаток при делении произведения kl на n. Мы принимаем поэтому такое определение *умножения классов*:

$$C_k \cdot C_l = C_r, \ \text{где} \ kl = nq + r, \ 0 \leqslant r < n. \qquad (3)$$

Система (1) *классов целых чисел, сравнимых между собой по модулю* n, *будет кольцом по отношению к операциям, определенным условиями* (2) *и* (3). В самом деле, справедливость требований I—V из определения кольца без труда устанавливается непосредственной проверкой, но вытекает также из справедливости этих требований в кольце целых чисел и той связи между операциями над целыми числами и операциями над классами, которая указана выше. Роль нуля играет, очевидно, класс C_0, состоящий из чисел, нацело делящихся на n. Противоположным для класса C_k, $k=1, 2, \ldots, n-1$, будет класс C_{n-k}. В системе классов (1) можно

определить, следовательно, вычитание, т. е. эта система удовлетворяет всем требованиям, входящим в определение кольца. Условимся обозначать полученное кольцо через Z_n.

Если число n составное, то кольцо Z_n обладает делителями нуля и поэтому, как будет показано ниже, не может быть полем. В самом деле, если $n=kl$, где $1<k<n$, $1<l<n$, то классы C_k и C_l отличны от нулевого класса C_0, но на основании определения умножения классов (см. (3)) $C_k \cdot C_l = C_0$.

Если же число n простое, то кольцо Z_n будет полем.

В самом деле, пусть даны классы C_k и C_m, причем $C_k \neq C_0$, т. е. $1 \leqslant k \leqslant n-1$. Нужно показать, что можно разделить C_m на C_k, т. е. найти такой класс C_l, что $C_k \cdot C_l = C_m$. Если $C_m = C_0$, то и $C_l = C_0$. Если же $C_m \neq C_0$, то рассмотрим систему чисел

$$k,\ 2k,\ 3k,\ \ldots,\ (n-1)\,k. \tag{4}$$

Все эти числа лежат вне нулевого класса C_0, так как произведение двух натуральных чисел, меньших простого числа n, не может на n делиться. Далее, никакие два числа sk и tk из системы (4), $s<t$, не могут лежать в одном классе, так как тогда их разность

$$tk - sk = (t-s)\,k$$

делилась бы на n, что снова противоречит простоте числа n. Таким образом, в каждом ненулевом классе лежит ровно одно число из системы (4). В частности, в классе C_m лежит число lk, где $1 \leqslant l \leqslant n-1$, т. е. $C_l \cdot C_k = C_m$, а тогда класс C_l и будет искомым частным от деления C_m на C_k.

Мы получили, таким образом, бесконечно много различных конечных полей: поле Z_2, состоящее всего из двух элементов, а также поля Z_3, Z_5, Z_7, Z_{11} и т. д.

Переходим к рассмотрению некоторых свойств полей, вытекающих из существования деления. Эти свойства аналогичны свойствам колец, основанным на существовании вычитания, и доказываются такими же рассуждениями, поэтому проведение доказательств предоставляется читателю.

Всякое поле P обладает однозначно определенным элементом, произведение которого на любой элемент a этого поля равно a. Этот элемент, совпадающий с равными между собою частными $\frac{a}{a}$ для всех a, отличных от нуля, называется *единицей* поля P и обозначается символом 1. Таким образом,

$$a \cdot 1 = a \quad \text{для всех } a \text{ из } P.$$

Во всяком поле для любого элемента a, отличного от нуля, существует однозначно определенный обратный элемент a^{-1}, удовлетворяющий равенству

$$a \cdot a^{-1} = 1,$$

а именно, $a^{-1}=\frac{1}{a}$. Очевидно, что $(a^{-1})^{-1}=a$. Частное $\frac{b}{a}$ можно записать теперь в виде

$$\frac{b}{a}=b\cdot a^{-1}.$$

Для любого элемента a, отличного от нуля, и любого целого положительного числа n имеет место равенство

$$(a^{-1})^n=(a^n)^{-1}.$$

Обозначая эти равные между собою элементы через a^{-n}, мы приходим к *отрицательным степеням* элемента поля, для которых сохраняются обычные правила оперирования. Положим, наконец, $a^0=1$ для всех a.

Существование единицы не является характерным свойством полей: единицей обладает, например, кольцо целых чисел. Вместе с тем, пример кольца четных чисел показывает, что не все кольца обладают единицей. С другой стороны, *всякое кольцо, обладающее единицей и содержащее обратный элемент для любого элемента, отличного от нуля, будет полем.* Действительно, в этом случае частным $\frac{b}{a}$, $a\neq 0$, будет служить произведение ba^{-1}. Единственность этого частного доказывается без затруднений.

Заметим, что *никакое поле не содержит делителей нуля.* Действительно, пусть $ab=0$, но $a\neq 0$. Умножая обе части равенства на элемент a^{-1}, мы получим слева $(a^{-1}a)b=1\cdot b=b$, а справа $a^{-1}\cdot 0=0$, т. е. $b=0$. Отсюда следует, что *во всяком поле любое равенство можно сократить на общий множитель, отличный от нуля.* В самом деле, если $ac=bc$ и $c\neq 0$, то $(a-b)c=0$, откуда $a-b=0$, т. е. $a=b$.

Из определения частного $\frac{a}{b}$ (где $b\neq 0$) и доказанной выше возможности записывать его в виде произведения ab^{-1} без труда может быть выведено, что *во всяком поле сохраняются все обычные правила обращения с дробями,* а именно:

$$\frac{a}{b}=\frac{c}{d} \text{ тогда и только тогда, если } ad=bc;$$

$$\frac{a}{b}\pm\frac{c}{d}=\frac{ad\pm bc}{bd};$$

$$\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd};$$

$$\frac{-a}{b}=-\frac{a}{b}.$$

Характеристика поля. Не все свойства числовых полей сохраняются в случае произвольного поля. Так, складывая число 1 само с собою несколько раз, т. е. беря любое целое положительное

кратное единицы, мы никогда не получим нуля, и вообще все эти кратные, т. е. все натуральные числа, отличны друг от друга. Если же мы будем брать целые кратные единицы в каком-либо конечном поле, то среди них непременно будут равные, так как это поле обладает лишь конечным числом различных элементов. Если все целые кратные единицы поля P являются различными элементами поля P, т. е. $k\cdot 1 \neq l\cdot 1$ при $k \neq l$, то говорят, что поле P имеет *характеристику нуль*; таковы, например, все числовые поля. Если же существуют такие целые числа k и l, что $k > l$, но в P имеет место равенство $k\cdot 1 = l\cdot 1$, то $(k-l)\cdot 1 = 0$, т. е. в P существует такое положительное кратное единицы, которое оказывается равным нулю. В этом случае P называется полем *конечной характеристики*, а именно *характеристики* p, если p есть тот первый положительный коэффициент, с которым единица поля P обращается в нуль. Примерами полей конечной характеристики служат все конечные поля; существуют, впрочем, и бесконечные поля, имеющие конечную характеристику.

Если поле P имеет характеристику p, то число p будет простым.

Действительно, из равенства $p = st$, где $s < p$, $t < p$, вытекало бы равенство $(s\cdot 1)(t\cdot 1) = p\cdot 1 = 0$, т. е., так как в поле не может быть делителей нуля, или $s\cdot 1 = 0$, или $t\cdot 1 = 0$, что, однако, противоречит определению характеристики как наименьшего положительного коэффициента, обращающего единицу поля в нуль.

Если характеристика поля P равна p, то для любого элемента a из этого поля имеет место равенство $pa = 0$. Если же характеристика поля P равна 0 и a — элемент этого поля, n — целое число, то из $a \neq 0$ и $n \neq 0$ следует $na \neq 0$.

Действительно, в первом случае элемент pa, т. е. сумму p слагаемых, равных a, можно, вынося a за скобки, представить в виде

$$pa = a(p\cdot 1) = a\cdot 0 = 0.$$

Во втором случае из равенства $na = 0$, т. е. $a(n\cdot 1) = 0$, следовало бы при $a \neq 0$ равенство $n\cdot 1 = 0$, т. е., так как характеристика поля равна нулю, $n = 0$.

Подполя, расширения. Пусть в поле P некоторая часть его элементов, составляющая множество P', сама оказывается полем по отношению к тем операциям, которые определены в поле P, т. е. для любых двух элементов a, b из P' содержащиеся в поле P элементы $a+b$, ab, $a-b$ и, при $b \neq 0$, $\frac{a}{b}$ принадлежат к P' (законы I—V, выполняясь в P, будут, конечно, выполняться и в P'). Тогда P' называется *подполем* поля P, а P — *расширением* поля P'. Понятно, что нуль и единица поля P будут содержаться также в P' и служить для P' нулем и единицей. Так, поле рациональных чисел

является подполем поля действительных чисел; все числовые поля будут подполями поля комплексных чисел.

Пусть в поле P даны подполе P' и элемент c, лежащий вне P', и пусть мы нашли минимальное подполе P'' поля P, содержащее и P', и c. Такое минимальное подполе может быть только одно, так как если бы P''' было еще одно подполе с этими свойствами, то пересечение подполей P'' и P''' (т. е. совокупность элементов, общих обоим подполям) содержало бы P' и элемент c и вместе с любыми двумя своими элементами содержало бы их сумму (эта сумма должна содержаться и в P'', и в P''', а потому и в их пересечении), а также их произведение, разность и частное; иными словами, это пересечение само было бы подполем, в противоречие с минимальностью подполя P''. Мы будем говорить, что поле P'' *получено присоединением к полю P' элемента c*, и употреблять запись $P''=P'(c)$.

Понятно, что поле $P'(c)$ содержит, помимо элемента c и всех элементов поля P', также все элементы, которые получаются из них при помощи сложения, умножения, вычитания и деления. В качестве примера укажем на рассматривавшееся в § 43 расширение поля рациональных чисел, состоящее из чисел вида $a+b\sqrt{2}$ с рациональными a, b: это расширение получается присоединением к полю рациональных чисел числа $\sqrt{2}$.

§ 46*. Изоморфизм колец (полей). Единственность поля комплексных чисел

В теории колец большую роль играет понятие изоморфизма. Именно, кольца L и L' называются *изоморфными*, если между ними можно установить такое взаимно однозначное соответствие, при котором для любых элементов a, b из L и соответствующих им элементов a', b' из L' сумме $a+b$ соответствует сумма $a'+b'$, а произведению ab соответствует произведение $a'b'$.

Пусть между кольцами L и L' установлено изоморфное соответствие. *При этом соответствии нулю* 0 *кольца L соответствует нуль* $0'$ *кольца L'*. Действительно, пусть элементу 0 соответствует элемент c' из L'. Берем произвольный элемент a из L и соответствующий ему элемент a' из L'. Тогда элементу $a+0$ должен соответствовать элемент $a'+c'$; но $a+0=a$, поэтому $a'+c'=a'$, откуда $c'=0'$. Далее, *элементу $-a$ соответствует элемент $-a'$*. Действительно, пусть элементу $-a$ соответствует элемент d'. Тогда элементу $a+(-a)=0$ должен соответствовать элемент $a'+d'$, т. е. $a'+d'=0'$, откуда $d'=-a'$. Отсюда следует, что *разности элементов из L соответствует разность соответствующих элементов в L'*. Аналогичными рассуждениями можно показать, что если кольцо L обладает единицей, то образ этого элемента (т. е. элемент, соответствующий ему в L' при рассматриваемом изоморфизме) будет единицей кольца L', и если элемент a из L

обладает обратным элементом a^{-1}, то образом элемента a^{-1} в L' будет элемент, обратный к a'.

Отсюда следует, что *кольцо, изоморфное полю, само будет полем.* Легко видеть также, что свойство кольца не иметь делителей нуля также сохраняется при изоморфном соответствии. Вообще, изоморфные кольца могут отличаться друг от друга природой своих элементов, но они тождественны по своим алгебраическим свойствам; всякая теорема, доказанная относительно некоторого кольца, будет справедливой для всех колец, с ним изоморфных, если только в доказательстве теоремы использовались лишь свойства операций, а не индивидуальные свойства элементов этого кольца. По этой причине *мы не будем считать изоморфные кольца или поля различными*; они будут для нас лишь разными экземплярами одного и того же кольца или поля.

Применим это понятие к вопросу о построении поля комплексных чисел. Изложенная в § 17 конструкция поля комплексных чисел, основанная на использовании точек плоскости, не является единственно возможной. Вместо точек можно было бы взять отрезки (векторы) на плоскости, выходящие из начала координат, и, задавая эти векторы их компонентами a, b на осях координат, определить сложение и умножение векторов при помощи тех же самых формул (2) и (3) из § 17, как и в случае точек плоскости. Можно было бы, далее, вообще отказаться от привлечения геометрического материала; замечая, что и точки плоскости, и векторы на плоскости задаются упорядоченными парами действительных чисел (a, b), можно просто взять совокупность всех таких пар и в ней ввести сложение и умножение по формулам (2) и (3) из указанного параграфа.

На самом деле все эти поля оказались бы по своим алгебраическим свойствам неразличимыми, как показывает следующая теорема:

Все расширения поля действительных чисел D, полученные присоединением к полю D корня уравнения

$$x^2+1=0, \tag{1}$$

изоморфны между собой.

Пусть, в самом деле, дано какое-либо поле P, являющееся расширением поля D и содержащее элемент, удовлетворяющий уравнению (1). Выбор обозначения для этого элемента находится в нашем распоряжении, и мы употребим для этой цели букву i. Таким образом, имеет место равенство $i^2+1=0$ (откуда $i^2=-1$), где возведение в степень и сложение нужно понимать в смысле операций, определенных в поле P. Мы хотим найти сейчас поле $D(i)$, получающееся присоединением к полю D элемента i, т. е. найти минимальное подполе поля P, содержащее и поле D, и элемент i.

Рассмотрим для этой цели все те элементы α поля P, которые можно записать в виде

$$\alpha=a+bi, \tag{2}$$

где a и b — произвольные действительные числа, а произведение числа b на элемент i и сумму числа a с этим произведением следует понимать в смысле операций, определенных в поле P. Никакой элемент α поля P не может обладать двумя различными записями такого вида: из

$$\alpha = a + bi = \bar{a} + \bar{b}i$$

и $b \neq \bar{b}$ следовало бы

$$i = \frac{\bar{a} - a}{b - \bar{b}},$$

т. е. i оказалось бы действительным числом; если же $b = \bar{b}$, то и $a = \bar{a}$. К числу элементов поля P, записываемых в виде (2), принадлежат, в частности, все действительные числа (случай $b = 0$), а также сам элемент i (случай $a = 0$, $b = 1$).

Покажем, что ***совокупность всех элементов вида*** (2) ***составляет подполе поля*** P; это и будет тогда искомым полем $D(i)$. Пусть нам даны элементы $\alpha = a + bi$ и $\beta = c + di$. Тогда, используя коммутативность и ассоциативность сложения и закон дистрибутивности, имеющие место в поле P, получаем:

$$\alpha + \beta = (a + bi) + (c + di) = (a + c) + (bi + di),$$

откуда

$$\alpha + \beta = (a + c) + (b + d)\,i, \tag{3}$$

т. е. эта сумма снова принадлежит к рассматриваемому множеству элементов. Далее,

$$-\beta = (-c) + (-d)\,i,$$

так как, ввиду (3), тогда будет справедливо равенство $\beta + (-\beta) = 0 + 0i = 0$; поэтому

$$\alpha - \beta = \alpha + (-\beta) = (a - c) + (b - d)\,i, \tag{3'}$$

т. е. и вычитание не выводит нас за пределы рассматриваемого множества. Снова используя свойства I—V, имеющие место для операций в поле P (см. § 44), и опираясь на равенство $i^2 = -1$, мы получаем:

$$\alpha\beta = (a + bi)(c + di) = ac + adi + bci + bdi^2,$$

т. е.

$$\alpha\beta = (ac - bd) + (ad + bc)\,i; \tag{4}$$

таким образом, произведение двух любых элементов вида (2) снова будет элементом этого же вида. Предположим, наконец, что $\beta \neq 0$, т. е. хотя бы одно из чисел c, d отлично от нуля. Тогда будет также $c - di \neq 0$ и

$$(c + di)(c - di) = c^2 - (di)^2 = c^2 - d^2i^2 = c^2 + d^2,$$

причем $c^2+d^2 \neq 0$. Поэтому, используя отмечавшееся в предшествующем параграфе утверждение, что во всяком поле сохраняются все обычные правила обращения с дробями, а поэтому, в частности, дробь не меняется от умножения ее числителя и знаменателя на один и тот же отличный от нуля элемент, получаем:

$$\frac{\alpha}{\beta}=\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2},$$

т. е. элемент

$$\frac{\alpha}{\beta}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i \qquad (4')$$

снова имеет вид (2).

Покажем теперь, что *полученное нами подполе $D(i)$ поля P изоморфно тому полю из точек плоскости, которое было построено в* § 17. Сопоставляя элементу $a+bi$ поля $D(i)$ точку (a, b), мы получим ввиду доказанной единственности записи вида (2) для элементов поля $D(i)$ взаимно однозначное соответствие между элементами этого поля и всеми точками плоскости. При этом соответствии действительному числу a соответствует точка $(a, 0)$ ввиду равенства $a=a+0i$, а элементу $i=0+1\cdot i$ сопоставляется точка $(0, 1)$. С другой стороны, сравнивая формулы (3) и (4) настоящего параграфа с формулами (2) и (3) из § 17, мы получаем, что сумме и произведению элементов α и β поля $D(i)$ сопоставляются точки, являющиеся суммой и соответственно произведением точек, сопоставленных элементам α и β.

Этим, так как все поля, изоморфные некоторому данному полю, изоморфны между собой, заканчивается доказательство теоремы. Мы видим, в частности, что выбор в § 17 формул (2) и (3) для определения операций над точками не был случайным и не может быть изменен.

Помимо способов построения поля комплексных чисел, рассматривавшихся выше, существуют и многие другие. Укажем один из них, использующий сложение и умножение матриц.

Рассмотрим некоммутативное кольцо матриц второго порядка над полем действительных чисел. Очевидно, что скалярные матрицы

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

составляют в этом кольце подполе, изоморфное полю действительных чисел. Оказывается, однако, что *в кольце матриц второго порядка над полем действительных чисел можно найти также подполе, изоморфное полю комплексных чисел.* В самом деле, поставим в соответствие всякому комплексному числу $a+bi$ *матрицу*

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Этим путем все поле комплексных чисел отображается, притом взаимно однозначно, на часть кольца матриц второго порядка, причем из равенств

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix}$$

вытекает, что это отображение изоморфное, так как матрицы, стоящие в правых частях равенств, соответствуют комплексным числам $(a+c)+(b+d)\,i=(a+bi)+(c+di)$ и $(ac-bd)+(ad+bc)\,i=(a+bi)\,(c+di)$. В частности, роль мнимой единицы i играет матрица

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Полученный нами результат указывает на еще один возможный способ построения поля комплексных чисел, столь же удовлетворительный, как и те, которые рассматривались выше.

§ 47. Линейная алгебра и алгебра многочленов над произвольным полем

В тех из предшествующих глав книги, которые посвящены линейной алгебре, роль основного поля играло обычно поле действительных чисел. Без труда проверяется, однако, что очень многое из этих глав дословно переносится на случай произвольного основного поля.

Так, *для произвольного основного поля P остаются справедливыми изложенные в гл. 1 метод Гаусса для решения систем линейных уравнений, теория определителей и правило Крамера.* Лишь замечание о кососимметрических определителях, приведенное в конце § 4, требует предположения, что характеристика поля P отлична от двух. Впрочем, доказательство свойства 4 из этого же параграфа также теряет силу, если характеристика поля P равна двум, хотя само это свойство остается справедливым.

Полезно отметить также, что неоднократно высказывавшееся в гл. 1 утверждение о существовании у неопределенной системы линейных уравнений бесконечного множества различных решений сохраняет силу в случае любого бесконечного основного поля P, но перестает быть справедливым, если поле P конечно.

Далее, *полностью переносятся на случай произвольного основного поля изложенные в гл. 2 теория линейной зависимости векторов, теория ранга матрицы и общая теория систем линейных уравнений, а также алгебра матриц из гл. 3.*

Общая теория квадратичных форм, построенная в § 26, *переносится на случай любого основного поля P, характеристика которого отлична от двух.* Без этого ограничения, как легко показать, основная теорема этого параграфа уже перестает быть справедливой.

Пусть, например, $P=Z_2$, т. е. является полем, состоящим из двух элементов 0 и 1, причем $1+1=0$, откуда $-1=1$, и пусть над этим полем дана квадратичная форма $f=x_1x_2$. Если существует линейное преобразование

$$x_1=b_{11}y_1+b_{12}y_2,$$
$$x_2=b_{21}y_1+b_{22}y_2,$$

приводящее f к каноническому виду, то в равенстве

$$f=(b_{11}y_1+b_{12}y_2)(b_{21}y_1+b_{22}y_2)=b_{11}b_{21}y_1^2+(b_{11}b_{22}+b_{12}b_{21})\,y_1y_2+b_{12}b_{22}y_2^2$$

коэффициент $b_{11}b_{22}+b_{12}b_{21}$ при произведении y_1y_2 должен быть равен нулю. Этот коэффициент равен, однако, определителю взятого нами линейного преобразования, так как будет ли $b_{12}b_{21}=1$ или же $b_{12}b_{21}=0$, — в обоих случаях $b_{12}b_{21}=-b_{12}b_{21}$. Наше линейное преобразование оказалось вырожденным.

Дальнейшее содержание гл. 6 существенно относится к квадратичным формам с комплексными или действительными коэффициентами.

Наконец, *для случая произвольного основного поля P сохраняется вся построенная в гл. 7 теория линейных пространств и их линейных преобразований*. Впрочем, понятие характеристического корня связано с теорией многочленов над произвольным полем, о которой речь будет идти ниже. Заметим, что теорема из § 33 о связи между характеристическими корнями и собственными значениями примет теперь следующую формулировку: характеристические корни линейного преобразования φ, лежащие в основном поле P, и только они, служат собственными значениями этого преобразования.

Что же касается теории евклидовых пространств (гл. 8), то она существенно связана с полем действительных чисел.

На случай произвольного основного поля P могут быть перенесены и некоторые из изложенных выше разделов алгебры многочленов. Предварительно необходимо, однако, придать точный смысл понятию многочлена над произвольным полем.

Дело в том, что в § 20 указывались две точки зрения на понятие многочлена — формально-алгебраическая и теоретико-функциональная. Они обе могут быть перенесены на случай произвольного основного поля. Будучи, однако, равносильными для случая числовых полей (см. § 24) и, как легко проверить, для бесконечных полей вообще, *для конечных полей они уже перестают быть равносильными*.

Рассмотрим, например, введенное в § 45 поле Z_2, состоящее из двух элементов 0 и 1, причем $1+1=0$. Многочлены $x+1$ и x^2+1 с коэффициентами из этого поля являются различными, т. е. не удовлетворяют алгебраическому определению равенства многочленов. Вместе с тем, оба эти многочлена при $x=0$ получают значение 1, а при $x=1$ — значение 0, т. е. как «функции» от «переменного» x, принимающего значения в поле Z_2, они должны считаться равными. В поле Z_3, состоящем из трех элементов: 0, 1, 2, причем

$1+2=0$, в таком же положении находятся многочлены x^3+x+1 и $2x+1$. Такие примеры можно указать вообще для всех конечных полей.

Таким образом, в теории, относящейся к случаю произвольного поля P, невозможно принять теоретико-функциональную точку зрения на многочлены. Необходимо, следовательно, придать полную ясность формально-алгебраическому определению многочлена. С этой целью мы проведем такое построение кольца многочленов над произвольным полем P, которое не использует с самого начала обычной записи многочленов через «неизвестное» x.

Рассмотрим всевозможные упорядоченные конечные системы элементов поля P, имеющие вид

$$(a_0,\ a_1,\ \ldots,\ a_{n-1},\ a_n), \tag{1}$$

причем n произвольно, $n \geqslant 0$, но при $n>0$ должно быть $a_n \neq 0$. Определяя для систем вида (1) сложение и умножение в соответствии с формулами (3) и (4) § 20, мы превратим совокупность этих систем в коммутативное кольцо; доказательства необходимых для этого свойств дословно повторяют то, что делалось в § 20 для числовых многочленов.

В построенном нами кольце системы вида (a) (случай $n=0$) составляют подполе, изоморфное полю P. Это позволяет отождествить такие системы с соответствующими элементами a поля P, т. е. положить

$$(a)=a \text{ для всех } a \text{ из } P. \tag{2}$$

С другой стороны, обозначим систему $(0,\ 1)$ буквой x,

$$x=(0,\ 1).$$

Тогда, применяя указанное выше определение умножения, мы получим, что $x^2=(0,\ 0,\ 1)$ и вообще

$$x^k=(\underbrace{0,\ 0,\ \ldots,0}_{k\text{ раз}},\ 1). \tag{3}$$

Используя теперь определения сложения и умножения упорядоченных систем, а также равенства (2) и (3), мы получим:

$$\begin{aligned}
&(a_0,\ a_1,\ a_2,\ \ldots,\ a_{n-1},\ a_n)=\\
&\quad=(a_0)+(0,\ a_1)+(0,\ 0,\ a_2)+\ldots\\
&\qquad\ldots+(\underbrace{0,\ 0,\ \ldots,\ 0}_{n-1\text{ раз}},\ a_{n-1})+(\underbrace{0,\ 0,\ \ldots,\ 0}_{n\text{ раз}},\ a_n)=\\
&\quad=(a_0)+(a_1)(0,\ 1)+(a_2)(0,\ 0,\ 1)+\ldots\\
&\qquad\ldots+(a_{n-1})(\underbrace{0,\ 0,\ \ldots,\ 0}_{n-1\text{ раз}},\ 1)+(a_n)(\underbrace{0,\ 0,\ \ldots,\ 0}_{n\text{ раз}},\ 1)=\\
&\qquad\qquad=a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}+a_nx^n.
\end{aligned}$$

Таким образом, всякая упорядоченная система вида (1) может быть записана в виде многочлена относительно x с коэффициентами из поля P, причем эта запись будет, очевидно, однозначной. Опираясь, наконец, на уже доказанную коммутативность сложения, можно перейти к записи по убывающим степеням x.

Мы построим, следовательно, коммутативное кольцо, которое естественно назвать *кольцом многочленов от неизвестного x над полем P*. Это кольцо обозначается символом $P[x]$.

В кольце $P[x]$ содержится само поле P, как уже было показано выше. Далее, как и в случае колец многочленов над числовыми полями (см. § 20), *кольцо $P[x]$ обладает единицей, не содержит делителей нуля и не является полем.*

Если поле P содержится в большем поле $\overline{P}$, то кольцо $P[x]$ будет подкольцом кольца $\overline{P}[x]$: всякий многочлен с коэффициентами из P можно считать, понятно, многочленом и над полем $\overline{P}$, а сумма и произведение многочленов зависят только от их коэффициентов и поэтому не меняются при переходе к большему полю.

Для того чтобы лучше представить себе истинный объем понятия «кольцо многочленов над полем P», посмотрим на него еще с одной стороны.

Пусть поле P содержится в качестве подкольца в некотором коммутативном кольце L. Элемент α кольца L называется *алгебраическим над полем P*, если существует такое уравнение n-й степени, $n \geqslant 1$, с коэффициентами из поля P, которому элемент α удовлетворяет; если же такого уравнения не существует, то элемент α называется *трансцендентным над полем P*. Понятно, что элемент x кольца $P[x]$ трансцендентен над полем P.

Справедлива следующая теорема:

Если элемент α кольца L трансцендентен над полем P, то подкольцо L', полученное присоединением элемента α к полю P (т. е. минимальное подкольцо кольца L, содержащее поле P и элемент α), *изоморфно кольцу многочленов $P[x]$*.

В самом деле, всякий элемент β кольца L, который может быть записан в виде

$$\beta = a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n, \qquad n \geqslant 0, \tag{4}$$

с коэффициентами $a_0, a_1, \ldots, a_{n-1}, a_n$ из поля P, будет содержаться в подкольце L'. Элемент β не может обладать двумя различными записями вида (4), так как, вычитая из одной записи другую, мы получили бы, что существует уравнение над полем P, удовлетворяемое элементом α, в противоречие с трансцендентностью этого элемента. Складывая элементы вида (4) по правилам сложения в кольце L, можно, понятно, складывать коэффициенты при одинаковых степенях α; это совпадает, однако, с правилом сложения многочленов. С другой стороны, перемножая элементы вида (4) по правилам

умножения в кольце L, мы можем, пользуясь законом дистрибутивности, совершить почленное перемножение, а затем собрать подобные члены; это приводит, очевидно, к известному нам правилу умножения многочленов. Этим доказано, что элементы вида (4) составляют в кольце L подкольцо, содержащее поле P и элемент α, т. е. совпадающее с L', и что это подкольцо изоморфно кольцу многочленов $P[x]$.

Мы видим, что сделанный выше выбор определений для операций над многочленами не был случайным: он вполне определяется тем, что элемент x кольца $P[x]$ должен быть трансцендентным над полем P.

Заметим, что при построении кольца многочленов $P[x]$ мы нигде не использовали деления элементов поля P и лишь один раз, а именно, при доказательстве утверждения о степени произведения многочленов, должны были бы сослаться на отсутствие в поле P делителей нуля. Можно, следовательно, взять произвольное коммутативное кольцо L и, повторяя проведенное выше построение, получить *кольцо многочленов* $L[x]$ *над кольцом* L; если при этом кольцо L не содержит делителей нуля, то степень произведения многочленов будет равна сумме степеней сомножителей и поэтому кольцо многочленов $L[x]$ также не будет содержать делителей нуля.

Возвращаясь к многочленам с коэффициентами из произвольного поля P, заметим, что на этот случай переносится по существу вся теория делимости многочленов, изложенная в §§ 20—22 нашей книги. Именно, *в кольце* $P[x]$ *имеет место алгоритм деления с остатком*, причем и частное, и остаток сами будут принадлежать к кольцу $P[x]$. Далее, *в кольце* $P[x]$ *имеет смысл понятие делителя и сохраняются все его основные свойства.* При этом то обстоятельство, что алгоритм деления не выводит за пределы основного поля P, позволяет утверждать, что *свойство многочлена* $\varphi(x)$ *быть делителем для* $f(x)$ *не зависит от того, рассматриваем ли мы поле P или же его любое расширение.*

В кольце $P[x]$ *сохраняются также определение и все свойства наибольшего общего делителя, в том числе сохраняются алгоритм Евклида и теорема, доказанная в* § 21 *при помощи этого алгоритма.* Заметим, что так как алгоритм деления с остатком не зависит, как мы знаем, от того, какое поле выбрано в качестве основного, то можно утверждать, что *наибольший общий делитель двух данных многочленов также не зависит от того, рассматриваем ли мы поле P или же его произвольное расширение* $\overline{P}$.

Наконец, *для многочленов над полем P сохраняет смысл понятие корня и остаются справедливыми основные свойства корней.* Сохраняется и теория кратных корней; впрочем, к этому вопросу мы вернемся еще раз в конце следующего параграфа.

Эти замечания позволят нам в дальнейшем при изучении многочленов над любым полем P ссылаться на § 20—22.

§ 48. Разложение многочленов на неприводимые множители

На основании теоремы о существовании корня в § 24 для полей комплексных и действительных чисел были доказаны существование и единственность разложения многочлена на неприводимые множители. Эти результаты являются частными случаями общих теорем, относящихся к многочленам над произвольным полем P. Настоящий параграф посвящается изложению этой общей теории, параллельной теории разложения целых чисел на простые множители.

Определим сначала те многочлены, которые играют в кольце многочленов такую же роль, какую в кольце целых чисел играют простые числа. Заранее подчеркнем, что в этом определении будет идти речь лишь о многочленах, степень которых больше или равна единице; это вполне соответствует тому, что при определении простых чисел и изучении разложений целых чисел на простые множители числа 1 и -1 исключаются из рассмотрения.

Пусть дан многочлен $f(x)$ степени n, $n \geqslant 1$, с коэффициентами из поля P. Ввиду свойства V из § 21 все многочлены нулевой степени будут служить делителями для $f(x)$. С другой стороны, по VII, делителями для $f(x)$ будут и все многочлены $cf(x)$, где c — отличный от нуля элемент из P, причем ими исчерпываются все делители многочлена $f(x)$, имеющие степень n. Что же касается делителей для $f(x)$, степень которых больше 0, но меньше n, то они могут в кольце $P[x]$ существовать, а могут и отсутствовать. В первом случае многочлен $f(x)$ называется *приводимым* в поле P (или над полем P), во втором случае — *неприводимым* в этом поле.

Вспоминая определение делителя, можно сказать, что *многочлен $f(x)$ степени n приводим в поле P, если он может быть разложен над этим полем (т. е. в кольце $P[x]$) в произведение двух множителей, степени которых меньше n*:

$$f(x) = \varphi(x)\psi(x), \tag{1}$$

и $f(x)$ неприводим в поле P, если в любом его разложении вида (1) *один из множителей имеет степень* 0, *другой — степень n*.

Следует обратить особое внимание на то обстоятельство, что о приводимости или неприводимости многочлена можно говорить лишь по отношению к данному полю P, так как многочлен, неприводимый в этом поле, может оказаться приводимым в некотором его расширении $\overline{P}$. Так, многочлен x^2-2 с целыми коэффициентами неприводим в поле рациональных чисел — он не может быть разложен в произведение двух множителей первой степени с рациональными коэффициентами. Однако в поле действительных чисел этот многочлен оказывается приводимым, как показывает равенство

$$x^2-2=(x-\sqrt{2})(x+\sqrt{2}).$$

Многочлен x^2+1 неприводим не только в поле рациональных чисел, но и в поле действительных чисел; он делается приводимым, однако, в поле комплексных чисел, так как

$$x^2+1=(x-i)(x+i).$$

Укажем некоторые основные свойства неприводимых многочленов, причем будем помнить, что речь идет о многочленах, неприводимых в поле P.

α) *Всякий многочлен первой степени неприводим.*

В самом деле, если бы этот многочлен был разложим в произведение множителей меньшей степени, то эти множители должны были бы иметь степень 0. Однако произведение любых многочленов нулевой степени снова будет многочленом нулевой степени, а не первой.

β) *Если многочлен $p(x)$ неприводим, то неприводимым будет и всякий многочлен $cp(x)$, где c — отличный от нуля элемент из P.*

Это свойство следует из свойств I и VII § 21. Оно позволит нам там, где это будет нужно, ограничиваться рассмотрением неприводимых многочленов, старшие коэффициенты которых равны единице.

γ) *Если $f(x)$ — произвольный, а $p(x)$ — неприводимый многочлен, то либо $f(x)$ делится на $p(x)$, либо же эти многочлены взаимно просты.*

Если $(f(x), p(x))=d(x)$, то $d(x)$, будучи делителем неприводимого многочлена $p(x)$, либо имеет степень 0, либо же есть многочлен вида $cp(x)$, $c\neq 0$. В первом случае $f(x)$ и $p(x)$ взаимно просты, во втором $f(x)$ делится на $p(x)$.

δ) *Если произведение многочленов $f(x)$ и $g(x)$ делится на неприводимый многочлен $p(x)$, то хотя бы один из этих множителей делится на $p(x)$.*

Действительно, если $f(x)$ не делится на $p(x)$, то, по γ), $f(x)$ и $p(x)$ взаимно просты, а тогда, по свойству б) из § 21, многочлен $g(x)$ должен делиться на $p(x)$.

Свойство δ) без труда распространяется на случай произведения любого конечного числа множителей.

Следующие две теоремы являются главной целью всего настоящего параграфа.

Всякий многочлен $f(x)$ из кольца $P[x]$, имеющий степень n, $n\geqslant 1$, разлагается в произведение неприводимых множителей.

Действительно, если многочлен $f(x)$ сам неприводим, то указанное произведение состоит всего из одного множителя. Если же он приводим, то может быть разложен в произведение множителей меньшей степени. Если среди этих множителей снова имеются приводимые, то производим их дальнейшее разложение на множители, и т. д. Этот процесс должен остановиться после конечного числа шагов, так как при любом разложении $f(x)$ на множители сумма

степеней этих множителей должна равняться n и поэтому число множителей, зависящих от x, не может превосходить n.

Разложение целых чисел на простые множители однозначно, если ограничиваться рассмотрением целых положительных чисел. Однако в кольце всех целых чисел однозначность имеет место лишь с точностью до знаков: так, $-6=2\cdot(-3)=(-2)\cdot 3$, $10=2\cdot 5=(-2)\cdot(-5)$ и т. д. Аналогичное положение имеет место и в кольце многочленов. Если

$$f(x)=p_1(x)\,p_2(x)\,\ldots\,p_s(x)$$

есть разложение многочлена $f(x)$ в произведение неприводимых множителей и если элементы $c_1, c_2, \ldots, c_s$ из поля P таковы, что их произведение равно 1, то

$$f(x)=[c_1p_1(x)]\cdot[c_2p_2(x)]\ldots[c_sp_s(x)]$$

также будет, ввиду β), разложением $f(x)$ в произведение неприводимых множителей. Оказывается, что этим исчерпываются все разложения $f(x)$:

Если многочлен $f(x)$ из кольца $P[x]$ двумя способами разложен в произведение неприводимых множителей:

$$f(x)=p_1(x)\,p_2(x)\,\ldots\,p_s(x)=q_1(x)\,q_2(x)\,\ldots\,q_t(x), \tag{2}$$

то $s=t$ и, при соответствующей нумерации, имеют место равенства

$$q_i(x)=c_ip_i(x),\quad i=1,\ 2,\ \ldots,\ s, \tag{3}$$

где c_i — отличные от нуля элементы из поля P.

Эта теорема верна для многочленов первой степени, так как они неприводимы. Мы будем поэтому вести доказательство индукцией по степени многочлена, т. е. будем доказывать теорему для $f(x)$, предполагая, что для многочленов меньшей степени она уже доказана.

Так как $q_1(x)$ является делителем для $f(x)$, то, ввиду свойства δ) и равенства (2), $q_1(x)$ будет делителем хотя бы для одного из многочленов $p_i(x)$, например для $p_1(x)$. Так как, однако, многочлен $p_1(x)$ неприводим, а степень $q_1(x)$ больше нуля, то существует такой элемент c_1, что

$$q_1(x)=c_1p_1(x). \tag{4}$$

Подставляя это выражение $q_1(x)$ в (2) и сокращая на $p_1(x)$ (что законно, так как в кольце $P[x]$ нет делителей нуля), мы получим равенство

$$p_2(x)\,p_3(x)\,\ldots\,p_s(x)=[c_1q_2(x)]\,q_3(x)\,\ldots\,q_t(x).$$

Так как степень многочлена, равного этим произведениям, меньше степени $f(x)$, то уже доказано, что $s-1=t-1$, откуда $s=t$, и что существуют такие элементы $c_2', c_3, \ldots, c_s$, что $c_2'\,p_2(x)=c_1q_2(x)$,

откуда $q_2(x) = (c_1^{-1} c_2') p_2(x)$, и $c_i p_i(x) = q_i(x)$, $i = 3, \ldots, s$. Полагая $c_1^{-1} c_2' = c_2$ и учитывая (4), мы полностью получим равенства (3).

Доказанной сейчас теореме можно дать такую более короткую формулировку: *всякий многочлен разлагается на неприводимые множители однозначно с точностью до множителей нулевой степени.*

Всегда можно рассматривать, впрочем, разложение следующего специального вида, *которое будет для каждого многочлена уже вполне однозначным*: берем любое разложение многочлена $f(x)$ на неприводимые множители и из каждого из этих множителей выносим за скобки старший коэффициент. Мы получим разложение

$$f(x) = a_0 p_1(x) p_2(x) \ldots p_s(x), \tag{5}$$

где все $p_i(x)$, $i = 1, 2, \ldots, s$, являются неприводимыми многочленами со старшими коэффициентами, равными единице. Множитель a_0 будет равен старшему коэффициенту многочлена $f(x)$, как легко доказать, выполнив перемножение в правой части равенства (5).

Неприводимые множители, входящие в разложение (5), не обязаны быть все различными. Если неприводимый многочлен $p(x)$ встречается в разложении (5) несколько раз, то он называется *кратным множителем* для $f(x)$, а именно *k-кратным* (в частности двукратным, трехкратным и т. д.), если в разложении (5) содержится ровно k множителей, равных $p(x)$. Если же множитель $p(x)$ входит в (5) лишь один раз, то он называется *простым* (или *однократным*) *множителем* для $f(x)$.

Если в разложении (5) множители $p_1(x), p_2(x), \ldots, p_l(x)$ отличны друг от друга, а всякий другой множитель равен одному из них, и если $p_i(x)$, $i = 1, 2, \ldots, l$, является k_i-кратным множителем многочлена $f(x)$, то разложение (5) можно переписать в следующем виде:

$$f(x) = a_0 p_1^{k_1}(x) p_2^{k_2}(x) \ldots p_l^{k_l}(x). \tag{6}$$

Именно этой записью мы будем дальше обычно пользоваться, не оговаривая особо, что показатели равны кратностям соответствующих множителей, т. е. что $p_i(x) \neq p_j(x)$ при $i \neq j$.

Если даны разложения многочленов $f(x)$ и $g(x)$ на неприводимые множители, то наибольший общий делитель $d(x)$ этих многочленов равен произведению множителей, входящих одновременно в оба разложения, причем каждый множитель берется в степени, равной меньшей из его кратностей в обоих данных многочленах.

Действительно, указанное произведение будет делителем для каждого из многочленов $f(x)$, $g(x)$, а поэтому и для $d(x)$. Если бы это произведение было отличным от $d(x)$, то в разложении $d(x)$ на неприводимые множители либо содержался бы множитель, который не входит в разложение хотя бы одного из многочленов $f(x)$ и $g(x)$,

что невозможно, либо же один из множителей имел бы бо́льшую степень, чем он имеет в разложении одного из многочленов $f(x)$ и $g(x)$, что снова невозможно.

Эта теорема аналогична тому правилу, по которому разыскивается обычно наибольший общий делитель целых чисел. Она не может заменить, однако, в случае многочленов алгоритм Евклида. Действительно, так как простых чисел, меньших данного целого положительного числа, лишь конечное число, то разложение целого числа на простые множители достигается конечным числом проб. Это уже не имеет места в кольце многочленов над бесконечным основным полем, и в общем случае нельзя дать способа для практического разложения многочленов на неприводимые множители. Больше того, даже решение вопроса, является ли многочлен $f(x)$ неприводимым в данном поле P, оказывается в общем случае весьма трудным. Так, описание всех неприводимых многочленов для случая полей комплексных и действительных чисел было получено в § 24 в качестве следствия из очень глубокой теоремы о существовании корня. Что же касается поля рациональных чисел, то о многочленах, неприводимых над этим полем, в § 56 будут сделаны лишь некоторые высказывания частного характера.

Мы показали, что в кольце многочленов, как и в кольце целых чисел, имеет место разложение на «простые» (неприводимые) множители и что это разложение в некотором смысле однозначно. Возникает вопрос, можно ли перенести эти результаты на более широкие классы колец. Мы ограничимся при этом случаем таких коммутативных колец, которые обладают единицей и не содержат делителей нуля.

Назовем *делителем единицы* такой элемент a кольца, для которого в этом кольце существует обратный элемент a^{-1},

$$aa^{-1}=1.$$

В кольце целых чисел это будут числа 1 и -1, в кольце многочленов $P[x]$ — все многочлены нулевой степени, т. е. отличные от нуля числа из поля P. Элемент c, отличный от нуля и не являющийся делителем единицы, назовем *простым* элементом кольца, если во всяком его разложении в произведение двух множителей, $c=ab$, один из этих множителей непременно является делителем единицы. В кольце целых чисел простыми элементами будут простые числа, в кольце многочленов — неприводимые многочлены.

Будет ли всякий элемент рассматриваемого кольца, отличный от нуля и не являющийся делителем единицы, разлагаться в произведение простых множителей? Если да, то будет ли такое разложение однозначным? Последнее нужно понимать в таком смысле: если

$$a=p_1p_2\ldots p_k=q_1q_2\ldots q_l$$

— два разложения элемента a на простые множители, то $k=l$ и (возможно, после изменения нумерации)

$$q_i=p_ic_i,\quad i=1, 2, \ldots, k,$$

где c_i — делитель единицы.

Оказывается, что в общем случае на оба вопроса должен быть дан отрицательный ответ. Мы ограничимся одним примером, а именно, укажем кольцо, в котором разложение на простые множители хотя и возможно, но не является однозначным.

Рассмотрим комплексные числа вида

$$\alpha = a + b\sqrt{-3}, \tag{7}$$

где a и b — целые числа. Все такие числа составляют кольцо без делителей нуля, содержащее единицу; действительно,

$$(a + b\sqrt{-3})(c + d\sqrt{-3}) = (ac - 3bd) + (bc + ad)\sqrt{-3}. \tag{8}$$

Назовем *нормой* числа $\alpha = a + b\sqrt{-3}$ целое положительное число

$$N(\alpha) = a^2 + 3b^2.$$

Ввиду (8) *норма произведения равна произведению норм,*

$$N(\alpha\beta) = N(\alpha)N(\beta). \tag{9}$$

Действительно,

$$(ac - 3bd)^2 + 3(bc + ad)^2 = a^2c^2 + 9b^2d^2 + 3b^2c^2 + 3a^2d^2 = (a^2 + 3b^2)(c^2 + 3d^2).$$

Если число α является в нашем кольце делителем единицы, т. е. число α^{-1} также имеет вид (7), то, по (9),

$$N(\alpha)\cdot N(\alpha^{-1}) = N(\alpha\alpha^{-1}) = N(1) = 1,$$

а поэтому $N(\alpha) = 1$, так как числа $N(\alpha)$ и $N(\alpha^{-1})$ — целые и положительные. Если $\alpha = a + b\sqrt{-3}$, то из $N(\alpha) = 1$ следует

$$N(\alpha) = a^2 + 3b^2 = 1;$$

это возможно, однако, лишь при $b = 0$, $a = \pm 1$. Таким образом, *в нашем кольце, как и в кольце целых чисел, делителями единицы будут лишь числа* 1 *и* -1 *и лишь эти числа имеют норму, равную единице.*

Равенство (9) для нормы произведения переносится, понятно, на случай любого конечного числа множителей. Отсюда легко вывести, что *всякое число* α *из нашего кольца может быть разложено в произведение конечного числа простых множителей;* проведение доказательства мы предоставим читателю.

Однозначность разложения на простые множители уже нельзя, однако, утверждать. Справедливы, например, равенства

$$4 = 2\cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}).$$

В нашем кольце нет других делителей единицы, кроме чисел 1 и -1, а поэтому число $1 + \sqrt{-3}$ (как и число $1 - \sqrt{-3}$) не может отличаться от числа 2 лишь на множитель, являющийся делителем единицы. Нам остается показать, что *каждое из чисел* 2, $1 + \sqrt{-3}$, $1 - \sqrt{-3}$ *будет в рассматриваемом кольце простым.* Действительно, норма каждого из этих трех чисел равна числу 4. Пусть α — любое из этих чисел и пусть

$$\alpha = \beta\gamma.$$

Тогда, по (9), возможен один из трех случаев:

1) $N(\beta) = 4$, $N(\gamma) = 1$; 2) $N(\beta) = 1$, $N(\gamma) = 4$; 3) $N(\beta) = N(\gamma) = 2$.

В первом случае число γ будет, как мы знаем, делителем единицы, во втором случае делителем единицы будет β. Что же касается третьего случая, то он вообще невозможен ввиду невозможности равенства

$$a^2 + 3b^2 = 2$$

при целых a и b.

Кратные множители. Хотя, как уже указано выше, мы не умеем разлагать многочлены на неприводимые множители, тем не менее существуют методы, позволяющие узнать, обладает ли данный многочлен кратными множителями, и в случае положительного ответа

дающие возможность свести изучение этого многочлена к изучению многочленов, уже не содержащих кратных множителей. Эти методы требуют, однако, наложения некоторых ограничений на основное поле. Именно, все дальнейшее содержание настоящего параграфа будет излагаться в предположении, что поле P имеет характеристику 0. Без этого ограничения теоремы о кратных множителях, которые будут доказаны ниже, уже теряют силу; вместе с тем, с точки зрения приложений, случай полей характеристики нуль является наиболее важным, так как сюда относятся, в частности, все числовые поля.

Заметим сначала, что на рассматриваемый случай переносятся и понятие производной многочлена, введенное в § 22 для многочленов с комплексными коэффициентами, и основные свойства этого понятия[1]). Докажем теперь следующую теорему:

Если $p(x)$ является k-кратным неприводимым множителем многочлена $f(x)$, $k \geqslant 1$, то он будет $(k-1)$-кратным множителем производной этого многочлена. В частности, простой множитель многочлена не входит в разложение производной.

В самом деле, пусть

$$f(x) = p^k(x)\, g(x), \tag{10}$$

причем $g(x)$ уже не делится на $p(x)$. Дифференцируя равенство (10), получаем:

$$f'(x) = p^k(x)\, g'(x) + kp^{k-1}(x)\, p'(x)\, g(x) = \\ = p^{k-1}(x)\, [p(x)\, g'(x) + kp'(x)\, g(x)].$$

Второе из слагаемых, стоящих в скобках, не делится на $p(x)$; действительно, $g(x)$ не делится на $p(x)$ по условию, $p'(x)$ имеет меньшую степень, т. е. также не делится на $p(x)$, а отсюда, ввиду неприводимости многочлена $p(x)$ и свойств δ) из настоящего параграфа и IX из § 21, следует наше утверждение. С другой стороны, первое слагаемое суммы, стоящей в квадратных скобках, делится на $p(x)$, а поэтому вся эта сумма не может делиться на $p(x)$, т. е. множитель $p(x)$ на самом деле входит в $f'(x)$ с кратностью $k-1$.

Из нашей теоремы и из указанного выше способа разыскания наибольшего общего делителя двух многочленов следует, что если дано разложение многочлена $f(x)$ на неприводимые множители:

$$f(x) = a_0 p_1^{k_1}(x)\, p_2^{k_2}(x) \,\ldots\, p_l^{k_l}(x), \tag{11}$$

то *наибольший общий делитель многочлена $f(x)$ и его производной обладает следующим разложением на неприводимые множители:*

$$(f(x),\ f'(x)) = p_1^{k_1-1}(x)\, p_2^{k_2-1}(x) \,\ldots\, p_l^{k_l-1}(x), \tag{12}$$

[1]) Для полей конечной характеристики теряет силу утверждение, что производная многочлена степени n имеет степень $n-1$.

где, понятно, множитель $p_i^{k_i-1}(x)$ следует при $k_i=1$ заменять единицей. В частности, *многочлен $f(x)$ тогда и только тогда не содержит кратных множителей, если он взаимно прост со своей производной.*

Мы научились, следовательно, отвечать на вопрос о существовании кратных множителей у данного многочлена. Больше того, так как ни производная многочлена, ни наибольший общий делитель двух многочленов не зависят от того, рассматриваем ли мы поле P или его любое расширение $\overline{P}$, то в качестве следствия из доказанного сейчас результата мы получаем:

Если многочлен $f(x)$ с коэффициентами из поля P характеристики нуль не имеет над этим полем кратных множителей, то у него не будет кратных множителей ни над каким расширением $\overline{P}$ поля P.

В частности, если $f(x)$ неприводим над P, а $\overline{P}$ — некоторое расширение поля P, то, хотя $f(x)$ уже может быть над $\overline{P}$ приводимым, однако заведомо не будет делиться на квадрат неприводимого (над $\overline{P}$) многочлена.

Выделение кратных множителей. Если дан многочлен $f(x)$ с разложением (11) и если через $d_1(x)$ мы обозначим наибольший общий делитель $f(x)$ и его производной $f'(x)$, то (12) будет разложением для $d_1(x)$. Деля (11) на (12), мы получим:

$$v_1(x)=\frac{f(x)}{d_1(x)}=a_0p_1(x)p_2(x)\ \dots\ p_l(x),$$

т. е. получим многочлен, не содержащий кратных множителей, причем всякий неприводимый множитель для $v_1(x)$ будет множителем и для $f(x)$. Этим разыскание неприводимых множителей для $f(x)$ сводится к разысканию их для многочлена $v_1(x)$, имеющего, вообще говоря, меньшую степень и, во всяком случае, содержащего лишь простые множители. Если эта задача для $v_1(x)$ будет решена, то останется определить лишь кратность найденных неприводимых множителей в $f(x)$, что достигается применением алгоритма деления.

Усложняя изложенный сейчас метод, можно сразу перейти к рассмотрению нескольких многочленов без кратных множителей, причем, найдя неприводимые множители этих многочленов, мы не только найдем все неприводимые множители для $f(x)$, но и будем знать их кратности.

Пусть (11) будет разложением $f(x)$ на неприводимые множители, причем наивысшая кратность множителей есть s, $s\geqslant 1$. Обозначим через $F_1(x)$ произведение всех однократных множителей многочлена $f(x)$, через $F_2(x)$ — произведение всех двукратных множителей, но взятых лишь по одному разу, и т. д., наконец, через $F_s(x)$ — произведение всех s-кратных множителей, также взятых лишь по одному разу; если при этом для некоторого j в $f(x)$ отсутствуют j-кратные множители, то полагаем $F_j(x)=1$. Тогда $f(x)$ будет делиться на k-ю степень многочлена $F_k(x)$, $k=1, 2, \dots, s$, и разложение (11) примет вид

$$f(x)=a_0F_1(x)F_2^2(x)F_3^3(x)\ \dots\ F_s^s(x),$$

а разложение (12) для $d_1(x)=(f(x),\ f'(x))$ перепишется в виде

$$d_1(x)=F_2(x)\,F_3^2(x)\,\dots\,F_s^{s-1}(x).$$

Обозначая через $d_2(x)$ наибольший общий делитель многочлена $d_1(x)$ и его производной и вообще через $d_k(x)$ наибольший общий делитель многочленов $d_{k-1}(x)$ и $d'_{k-1}(x)$, мы таким же путем получим:

$$d_2(x)=F_3(x)\,F_4^2(x)\,\dots\,F_s^{s-2}(x),$$
$$d_3(x)=F_4(x)\,F_5^2(x)\,\dots\,F_s^{s-3}(x),$$
$$\dots\dots\dots\dots\dots\dots$$
$$d_{s-1}(x)=F_s(x),$$
$$d_s(x)=1.$$

Отсюда

$$v_1(x)=\frac{f(x)}{d_1(x)}=a_0F_1(x)\,F_2(x)\,F_3(x)\,\dots\,F_s(x),$$
$$v_2(x)=\frac{d_1(x)}{d_2(x)}=F_2(x)\,F_3(x)\,\dots\,F_s(x),$$
$$v_3(x)=\frac{d_2(x)}{d_3(x)}=F_3(x)\,\dots\,F_s(x),$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$v_s(x)=\frac{d_{s-1}(x)}{d_s(x)}=F_s(x),$$

и поэтому, наконец,

$$F_1(x)=\frac{v_1(x)}{a_0v_2(x)},\quad F_2(x)=\frac{v_2(x)}{v_3(x)},\quad \dots,\quad F_s(x)=v_s(x).$$

Таким образом, пользуясь лишь приемами, не требующими знания неприводимых множителей многочлена $f(x)$, а именно взятием производной, алгоритмом Евклида и алгоритмом деления, мы можем найти многочлены $F_1(x)$, $F_2(x)$, ..., $F_s(x)$ без кратных множителей, причем всякий неприводимый множитель многочлена $F_k(x)$, $k=1, 2, \dots, s$, будет k-кратным для $f(x)$.

Изложенный здесь метод нельзя, понятно, считать методом для разложения многочлена на неприводимые множители, так как для случая $s=1$, т. е. для многочлена без кратных множителей, мы получим лишь $f(x)=F_1(x)$.

§ 49*. Теорема существования корня

Само собою разумеется, что доказанная в § 23 основная теорема о существовании для всякого числового многочлена корня в поле комплексных чисел не может быть перенесена на случай произвольного поля. В настоящем параграфе будет доказана теорема, в некоторой мере заменяющая в общей теории полей указанную основную теорему алгебры комплексных чисел.

Пусть дан многочлен $f(x)$ над полем P. Естественно возникает следующий вопрос: если многочлен $f(x)$ вообще не имеет корней в поле P, то существует ли такое расширение $\overline{P}$ поля P, в котором

для $f(x)$ уже найдется хотя бы один корень? При этом можно считать, что степень многочлена $f(x)$ больше единицы: для многочленов нулевой степени вопрос не имеет смысла, а всякий многочлен первой степени $ax+b$ обладает корнем $-\frac{b}{a}$ в самом поле P. С другой стороны, можно ограничиться, очевидно, случаем когда многочлен $f(x)$ неприводим: если он приводим над P, то корень любого из его неприводимых множителей будет служить корнем и для него самого.

Ответ на интересующий нас вопрос дает следующая **теорема существования корня**:

Для всякого многочлена $f(x)$, неприводимого над полем P, существует такое расширение этого поля, в котором содержится корень для $f(x)$. Все минимальные поля, содержащие поле P и какой-либо корень этого многочлена, изоморфны между собой.

Докажем сначала вторую половину этой теоремы.

Пусть дан неприводимый над P многочлен

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n, \tag{1}$$

причем $n\geqslant 2$, т. е. $f(x)$ не имеет корней в самом поле P. Предположим, что существует расширение $\overline{P}$ поля P, содержащее корень α для $f(x)$, и докажем следующую **лемму**, необходимую для дальнейшего, но представляющую и самостоятельный интерес:

Если лежащий в $\overline{P}$ корень α многочлена $f(x)$, неприводимого над P, служит корнем также для некоторого многочлена $g(x)$ из кольца $P[x]$, то $f(x)$ будет делителем для $g(x)$.

В самом деле, над полем $\overline{P}$ многочлены $f(x)$ и $g(x)$ обладают общим делителем $x-\alpha$ и поэтому не являются взаимно простыми. Свойство многочленов не быть взаимно простыми не зависит, однако, от выбора поля, поэтому можно перейти к полю P и применить свойство γ) из предшествующего параграфа.

Найдем теперь минимальное подполе $P(\alpha)$ поля $\overline{P}$, содержащее поле P и элемент α. К нему заведомо принадлежат все элементы вида

$$\beta=b_0+b_1\alpha+b_2\alpha^2+\ldots+b_{n-1}\alpha^{n-1}, \tag{2}$$

где $b_0, b_1, b_2, \ldots, b_{n-1}$ — элементы поля P. Никакой элемент поля $\overline{P}$ не может обладать двумя различными записями вида (2): если имеет место также равенство

$$\beta=c_0+c_1\alpha+c_2\alpha^2+\ldots+c_{n-1}\alpha^{n-1},$$

причем хотя бы при одном k $c_k\neq b_k$, то α будет корнем многочлена

$$g(x)=(b_0-c_0)+(b_1-c_1)x+(b_2-c_2)x^2+\ldots+(b_{n-1}-c_{n-1})x^{n-1},$$

что противоречит доказанной выше лемме, так как степень $g(x)$ меньше степени $f(x)$.

К числу элементов поля $\overline{P}$, имеющих вид (2), принадлежат все элементы поля P (при $b_1 = b_2 = \ldots = b_{n-1} = 0$), а также сам элемент α (при $b_1 = 1$, $b_0 = b_2 = \ldots = b_{n-1} = 0$). Докажем, что *элементы вида* (2) *составляют все искомое подполе* $P(\alpha)$. В самом деле, если даны элементы β (с записью (2)) и

$$\gamma = c_0 + c_1\alpha + c_2\alpha^2 + \ldots + c_{n-1}\alpha^{n-1},$$

то, на основании свойств операций в поле $\overline{P}$,

$$\beta \pm \gamma = (b_0 \pm c_0) + (b_1 \pm c_1)\alpha + (b_2 \pm c_2)\alpha^2 + \ldots + (b_{n-1} \pm c_{n-1})\alpha^{n-1},$$

т. е. сумма и разность двух любых элементов вида (2) снова будут элементами такого же вида.

Если мы перемножим β и γ, то получим выражение, содержащее α^n и более высокие степени α. Однако из (1) и равенства $f(\alpha) = 0$ вытекает, что α^n, а поэтому и α^{n+1}, α^{n+2} и т. д., можно выразить через меньшие степени элемента α. Наиболее простой способ разыскания выражения для $\beta\gamma$ состоит в следующем: пусть

$$\varphi(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1}, \quad \psi(x) = c_0 + c_1x + \ldots + c_{n-1}x^{n-1},$$

откуда $\varphi(\alpha) = \beta$, $\psi(\alpha) = \gamma$. Перемножим многочлены $\varphi(x)$ и $\psi(x)$ и разделим это произведение на $f(x)$; мы получим

$$\varphi(x)\psi(x) = f(x)q(x) + r(x), \tag{3}$$

где

$$r(x) = d_0 + d_1x + \ldots + d_{n-1}x^{n-1}.$$

Беря значения обеих частей равенства (3) при $x = \alpha$, мы получим:

$$\varphi(\alpha)\psi(\alpha) = f(\alpha)q(\alpha) + r(\alpha),$$

т. е., ввиду $f(\alpha) = 0$,

$$\beta\gamma = d_0 + d_1\alpha + \ldots + d_{n-1}\alpha^{n-1}.$$

Таким образом, произведение двух элементов вида (2) снова будет элементом такого же вида.

Покажем, наконец, что если элемент β имеет вид (2), причем $\beta \neq 0$, то элемент β^{-1}, существующий в поле $\overline{P}$, также может быть записан в виде (2). Для этого возьмем многочлен

$$\varphi(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1}$$

из кольца $P[x]$. Так как степень $\varphi(x)$ ниже степени $f(x)$, а многочлен $f(x)$ неприводим над P, то $\varphi(x)$ и $f(x)$ взаимно просты и поэтому, по §§ 21 и 47, в кольце $P[x]$ существуют такие многочлены $u(x)$ и $v(x)$, что

$$\varphi(x)u(x) + f(x)v(x) = 1;$$

можно считать при этом, что степень $u(x)$ меньше n:

$$u(x) = s_0 + s_1 x + \ldots + s_{n-1} x^{n-1}.$$

Отсюда, ввиду равенства $f(\alpha) = 0$, следует:

$$\varphi(\alpha) u(\alpha) = 1$$

и поэтому, ввиду равенства $\varphi(\alpha) = \beta$, мы получаем:

$$\beta^{-1} = u(\alpha) = s_0 + s_1 \alpha + \ldots + s_{n-1} \alpha^{n-1}.$$

Таким образом, совокупность элементов поля $\overline{P}$, имеющих вид (2), составляет подполе поля $\overline{P}$; это и будет искомое поле $P(\alpha)$. Так как мы видели, далее, что при разыскании суммы и произведения элементов β и γ вида (2) нужно знать лишь коэффициенты выражений этих элементов через степени α, то можно утверждать справедливость следующего результата: если существует, помимо $\overline{P}$, другое расширение $\overline{P'}$ поля P, также содержащее некоторый корень α' многочлена $f(x)$, и если $P(\alpha')$ есть минимальное подполе поля $\overline{P'}$, содержащее P и α', то *поля $P(\alpha)$ и $P(\alpha')$ будут изоморфными*, причем для получения изоморфного соответствия между ними нужно элементу β вида (2) из $P(\alpha)$ сопоставить элемент

$$\beta' = b_0 + b_1 \alpha' + b_2 \alpha'^2 + \ldots + b_{n-1} \alpha'^{n-1}$$

из $P(\alpha')$, имеющий те же коэффициенты. Этим доказана вторая половина теоремы.

Переходим к доказательству основной первой половины этой теоремы, причем изложенное выше подскажет нам пути для этого. Нам дан многочлен $f(x)$ степени $n \geqslant 2$, неприводимый над полем P, и нужно построить расширение поля P, содержащее корень для $f(x)$. Для этого возьмем все кольцо многочленов $P[x]$ и разобьем его на непересекающиеся классы, отнеся в один класс многочлены, дающие при делении на заданный нам многочлен $f(x)$ одинаковые остатки. Иными словами, многочлены $\varphi(x)$ и $\psi(x)$ относятся к одному классу, если их разность нацело делится на $f(x)$.

Условимся обозначать полученные классы буквами A, B, C и т. д. и следующим вполне естественным способом определим сумму и произведение классов. Возьмем любые два класса A и B, выберем в классе A некоторый многочлен $\varphi_1(x)$, в классе B — некоторый многочлен $\psi_1(x)$ и обозначим через $\chi_1(x)$ сумму этих многочленов,

$$\chi_1(x) = \varphi_1(x) + \psi_1(x),$$

а через $\theta_1(x)$ — их произведение,

$$\theta_1(x) = \varphi_1(x) \cdot \psi_1(x).$$

Выберем теперь в классе A любой другой многочлен $\varphi_2(x)$, в классе B — любой многочлен $\psi_2(x)$ и обозначим через $\chi_2(x)$ и $\theta_2(x)$ соответственно их сумму и произведение:

$$\chi_2(x)=\varphi_2(x)+\psi_2(x),$$
$$\theta_2(x)=\varphi_2(x)\cdot\psi_2(x).$$

По условию многочлены $\varphi_1(x)$ и $\varphi_2(x)$ лежат в одном классе A, а поэтому их разность $\varphi_1(x)-\varphi_2(x)$ нацело делится на $f(x)$; этим же свойством обладает и разность $\psi_1(x)-\psi_2(x)$. Отсюда следует, что разность

$$\chi_1(x)-\chi_2(x)=[\varphi_1(x)+\psi_1(x)]-[\varphi_2(x)+\psi_2(x)]=$$
$$=[\varphi_1(x)-\varphi_2(x)]+[\psi_1(x)-\psi_2(x)] \qquad (4)$$

также нацело делится на многочлен $f(x)$. Это же верно и для разности $\theta_1(x)-\theta_2(x)$, так как

$$\theta_1(x)-\theta_2(x)=\varphi_1(x)\psi_1(x)-\varphi_2(x)\psi_2(x)=$$
$$=\varphi_1(x)\psi_1(x)-\varphi_1(x)\psi_2(x)+\varphi_1(x)\psi_2(x)-\varphi_2(x)\psi_2(x)=$$
$$=\varphi_1(x)[\psi_1(x)-\psi_2(x)]+[\varphi_1(x)-\varphi_2(x)]\psi_2(x). \qquad (5)$$

Равенство (4) показывает, что многочлены $\chi_1(x)$ и $\chi_2(x)$ лежат в одном классе. Иными словами, сумма любого многочлена из класса A с любым многочленом из класса B принадлежит ко вполне определенному классу C, который не зависит от того, какие именно многочлены выбраны в качестве «представителей» в классах A и B; назовем этот класс C *суммой* классов A и B:

$$C=A+B.$$

Аналогично, ввиду (5), не зависит от выбора представителей в классах A и B и тот класс D, в котором лежит произведение любого многочлена из A на любой многочлен из B; этот класс назовем *произведением* классов A и B:

$$D=AB.$$

Покажем, что совокупность классов, на которые разбито нами кольцо многочленов $P[x]$, после указанного введения операций сложения и умножения превращается в поле. В самом деле, справедливость законов ассоциативности и коммутативности для обеих операций и закона дистрибутивности вытекает из справедливости этих законов в кольце $P[x]$, так как операции над классами сводятся на операции над многочленами, лежащими в этих классах. Роль нуля играет, очевидно, класс, составленный из многочленов, нацело делящихся на многочлен $f(x)$. Этот класс назовем *нулевым* и будем обозначать символом 0. Противоположным для класса A, составленного из многочленов, дающих

при делении на $f(x)$ остаток $\varphi(x)$, будет служить класс, составленный из многочленов, дающих при делении на $f(x)$ остаток $-\varphi(x)$. Отсюда вытекает, что в множестве классов выполнимо однозначное вычитание.

Для доказательства того, что в множестве классов выполнимо деление, нужно показать, что существует класс, играющий роль единицы, и что для всякого класса, отличного от нулевого, существует обратный класс. Единицей будет, очевидно, класс многочленов, дающих при делении на $f(x)$ остаток 1; этот класс назовем *единичным* и будем обозначать символом E.

Пусть теперь дан класс A, отличный от нулевого. Многочлен $\varphi(x)$, выбранный в классе A в качестве представителя, не будет, следовательно, нацело делиться на $f(x)$, и поэтому, ввиду неприводимости многочлена $f(x)$, эти два многочлена взаимно просты. В кольце $P[x]$ существуют, таким образом, многочлены $u(x)$ и $v(x)$, удовлетворяющие равенству

$$\varphi(x)\,u(x)+f(x)\,v(x)=1,$$

откуда

$$\varphi(x)\,u(x)=1-f(x)\,v(x). \qquad (6)$$

Правая часть равенства (6) при делении на $f(x)$ дает в остатке 1, т. е. принадлежит к единичному классу E. Если класс, к которому принадлежит многочлен $u(x)$, мы обозначим через B, то равенство (6) показывает, что

$$AB=E,$$

откуда $B=A^{-1}$. Этим доказано существование обратного класса для всякого ненулевого класса, т. е. закончено доказательство того, что классы составляют поле.

Обозначим это поле через $\overline{P}$ и покажем, что *оно является расширением поля P*. Всякому элементу a поля P соответствует класс, составленный из многочленов, дающих при делении на $f(x)$ остаток a; сам элемент a, рассматриваемый как многочлен нулевой степени, принадлежит к этому классу. Все классы этого специального вида составляют в поле $\overline{P}$ подполе, изоморфное полю P. Действительно, взаимная однозначность соответствия очевидна; с другой стороны, в этих классах можно выбрать в качестве представителей элементы поля P, а поэтому сумме (произведению) элементов из P будет соответствовать сумма (произведение) соответствующих классов. В дальнейшем мы имеем право, следовательно, не различать элементы поля P и соответствующие им классы.

Обозначим, наконец, через X класс, составленный из многочленов, дающих при делении на $f(x)$ остаток x. Этот класс является вполне определенным элементом поля $\overline{P}$, и мы хотим показать, что *он служит корнем для многочлена* $f(x)$. Пусть

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n.$$

Обозначим через A_i класс, соответствующий в указанном выше смысле элементу a_i поля P, $i=0, 1, \ldots, n$, и найдем, чему равен элемент

$$A_0X^n+A_1X^{n-1}+\ldots+A_{n-1}X+A_n \quad (7)$$

поля $\overline{P}$. Считая представителями классов A_i элементы $a_i, i=0,1,\ldots,n$, а представителем класса X — многочлен x и используя определение сложения и умножения классов, мы получаем, что в классе (7) содержится сам многочлен $f(x)$. Однако $f(x)$ нацело делится на самого себя, и поэтому класс (7) оказывается нулевым. Таким образом, заменяя в (7) классы A_i соответствующими им элементами a_i поля P, мы получаем, что в поле $\overline{P}$ имеет место равенство

$$a_0X^n+a_1X^{n-1}+\ldots+a_{n-1}X+a_n=0,$$

т. е. класс X действительно является корнем многочлена $f(x)$.

Этим заканчивается доказательство теоремы о существовании корня. Заметим, что, взяв за P поле действительных чисел и положив $f(x)=x^2+1$, мы получим еще один способ построения поля комплексных чисел.

Из теоремы о существовании корня могут быть выведены следствия, аналогичные тем, которые выводились в § 24 из основной теоремы алгебры комплексных чисел. Сначала сделаем одно замечание. Так как всякий линейный множитель $x-c$ многочлена $f(x)$ неприводим, то он должен входить в то единственное разложение на неприводимые множители, которым обладает $f(x)$.

Число линейных множителей в разложении $f(x)$ на неприводимые множители не может превосходить, однако, степени этого многочлена. Мы приходим к следующему результату:

Многочлен $f(x)$ степени n может иметь в поле P не более n корней, если даже каждый из корней считать столько раз, какова его кратность.

Назовем *полем разложения* для многочлена $f(x)$ степени n над полем P такое расширение Q поля P, в котором для $f(x)$ содержится n корней (считая кратные корни столько раз, какова их кратность). Над полем Q многочлен $f(x)$ будет раскладываться, следовательно, на линейные множители, причем никакое дальнейшее расширение поля Q уже не может привести к появлению новых корней для $f(x)$.

Для всякого многочлена $f(x)$ из кольца $P[x]$ существует над полем P поле разложения.

В самом деле, если многочлен $f(x)$ степени n, $n\geqslant 1$, имеет n корней в самом поле P, то P будет искомым полем разложения. Если же $f(x)$ не разлагается над P на линейные множители, то берем один из его нелинейных неприводимых множителей $\varphi(x)$ и, на основании теоремы о существовании корня, расширяем P до поля P', содержащего корень для $\varphi(x)$. Если над P' многочлен $f(x)$

все еще не разлагается на линейные множители, то снова расширяем поле, создавая корень еще для одного из оставшихся нелинейных неприводимых множителей. После конечного числа шагов мы придем, очевидно, к полю разложения для $f(x)$.

Понятно, что $f(x)$ может обладать многими различными полями разложения. Можно было бы доказать, что все минимальные поля, содержащие поле P и n корней многочлена $f(x)$ (где n — степень этого многочлена), изоморфны между собой. Мы не будем, однако, использовать этого утверждения и поэтому не приводим его доказательства.

Кратные корни. В предшествующем параграфе было доказано, что многочлен $f(x)$ над полем P характеристики 0 тогда и только тогда не имеет кратных множителей, если он взаимно прост со своей производной, а также было отмечено, что отсутствие у $f(x)$ кратных множителей над P влечет за собой отсутствие таких множителей над любым расширением $\overline{P}$ поля P. Применяя это к случаю когда $\overline{P}$ есть некоторое поле разложения для $f(x)$, и вспоминая определение кратного корня, мы приходим к следующему результату:

Если многочлен $f(x)$ над полем P характеристики 0 не имеет кратных корней в данном поле разложения, то он взаимно прост со своей производной $f'(x)$. Обратно, если $f(x)$ взаимно прост со своей производной, то он не имеет кратных корней ни в каком из своих полей разложения.

Отсюда, в частности, вытекает, что *многочлен $f(x)$, неприводимый над полем P характеристики 0, не может иметь кратных корней ни в каком расширении этого поля.* Для полей конечной характеристики это утверждение перестает быть справедливым — обстоятельство, играющее заметную роль в общей теории полей.

В заключение заметим, что *для случая произвольного поля сохраняются и формулы Вьета* (см. § 24); при этом корни многочлена берутся в некотором поле разложения этого многочлена.

§ 50*. Поле рациональных дробей

Теория рациональных дробей, изложенная в § 25, *полностью сохраняется и в случае произвольного основного поля.* Однако при переходе от поля действительных чисел к произвольному полю P взгляд на выражения $\frac{f(x)}{g(x)}$ как на функции переменного x должен быть отброшен, так как он, как мы знаем, неприменим уже к многочленам. Перед нами стоит задача определить, какой смысл нужно придать этим выражениям в том случае, когда коэффициенты принадлежат к произвольному полю P. Точнее, мы хотим построить поле, в котором содержалось бы кольцо многочленов $P[x]$, причем так, чтобы операции сложения и умножения, определенные в этом новом поле, в применении к многочленам совпадали бы

с операциями в кольце $P[x]$; короче, кольцо $P[x]$ должно быть **подкольцом** этого нового поля. С другой стороны, всякий элемент этого нового поля должен представляться (в смысле деления, определенного в этом поле) в виде частного двух многочленов. Такое поле для всякого P может быть построено, как будет сейчас показано; его обозначают $P(x)$ (неизвестное заключено в круглые скобки!) и называют *полем рациональных дробей* над полем P.

Предположим сначала, что кольцо $P[x]$ уже является подкольцом некоторого поля Q. Если $f(x)$ и $g(x)$ — произвольные многочлены из $P[x]$, причем $g(x) \neq 0$, то в поле Q существует однозначно определенный элемент, равный частному от деления $f(x)$ на $g(x)$. Обозначая этот элемент, как обычно в случае поля, через $\frac{f(x)}{g(x)}$, мы на основании определения частного можем написать равенство

$$f(x) = g(x) \cdot \frac{f(x)}{g(x)}, \tag{1}$$

где произведение нужно понимать в смысле умножения в поле Q. Может случиться, что некоторые частные $\frac{f(x)}{g(x)}$ и $\frac{\varphi(x)}{\psi(x)}$ являются одним и тем же элементом поля Q; условием для этого является обычное условие равенства дробей:

Тогда и только тогда $\frac{f(x)}{g(x)} = \frac{\varphi(x)}{\psi(x)}$, *если* $f(x)\psi(x) = \varphi(x)g(x)$.

Действительно, если $\frac{f(x)}{g(x)} = \frac{\varphi(x)}{\psi(x)} = \alpha$, то, по (1),

$$f(x) = g(x)\alpha, \quad \varphi(x) = \psi(x)\alpha,$$

откуда

$$f(x)\psi(x) = g(x)\psi(x)\alpha = g(x)\varphi(x).$$

Обратно, если $f(x)\psi(x) = g(x)\varphi(x) = u(x)$ в смысле умножения в кольце $P[x]$, то, переходя к полю Q, мы получаем равенства

$$\frac{f(x)}{g(x)} = \frac{u(x)}{g(x)\psi(x)} = \frac{\varphi(x)}{\psi(x)}.$$

Легко видеть, далее, что сумма и произведение любых элементов из Q, являющихся частными многочленов из $P[x]$, снова могут быть представлены в виде таких частных, причем справедливы обычные правила сложения и умножения дробей:

$$\frac{f(x)}{g(x)} + \frac{\varphi(x)}{\psi(x)} = \frac{f(x)\psi(x) + g(x)\varphi(x)}{g(x)\psi(x)}, \tag{2}$$

$$\frac{f(x)}{g(x)} \cdot \frac{\varphi(x)}{\psi(x)} = \frac{f(x)\cdot\varphi(x)}{g(x)\cdot\psi(x)}. \tag{3}$$

Действительно, умножая обе части каждого из этих равенств на произведение $g(x)\psi(x)$ и применяя (1), мы получим равенства, справедливые в кольце $P[x]$. Справедливость равенств (2) и (3)

следует теперь из того, что, благодаря отсутствию делителей нуля в поле Q, обе части каждого из полученных равенств можно сократить на отличный от нуля элемент $g(x)\psi(x)$, не нарушая равенств.

Эти предварительные замечания подсказывают нам тот путь, по которому мы должны пойти при построении поля $P(x)$. Пусть даны произвольное поле P и над ним кольцо многочленов $P[x]$. Всякой упорядоченной паре многочленов $f(x)$, $g(x)$, где $g(x)\neq 0$, мы ставим в соответствие символ $\frac{f(x)}{g(x)}$, называемый *рациональной дробью с числителем* $f(x)$ *и знаменателем* $g(x)$. Подчеркиваем, что это просто символ, соответствующий данной паре многочленов, так как деление многочленов в самом кольце $P[x]$, вообще говоря, невыполнимо, а ни в каком поле кольцо $P[x]$ пока еще не содержится; если даже $g(x)$ является делителем для $f(x)$, новый символ $\frac{f(x)}{g(x)}$ следует пока отличать от многочлена, получающегося в качестве частного при делении $f(x)$ на $g(x)$.

Назовем теперь рациональные дроби $\frac{f(x)}{g(x)}$ и $\frac{\varphi(x)}{\psi(x)}$ *равными*:

$$\frac{f(x)}{g(x)}=\frac{\varphi(x)}{\psi(x)}, \tag{4}$$

если в кольце $P[x]$ имеет место равенство $f(x)\psi(x)=g(x)\varphi(x)$. Очевидно, что всякая дробь равна самой себе, а также, что если одна дробь равна другой, то и вторая равна первой. Докажем транзитивность этого понятия равенства. Пусть даны равенства (4) и

$$\frac{\varphi(x)}{\psi(x)}=\frac{u(x)}{v(x)}. \tag{5}$$

Из равносильных им равенств в кольце $P[x]$

$$f(x)\psi(x)=g(x)\varphi(x), \qquad \varphi(x)v(x)=\psi(x)u(x)$$

вытекает

$$f(x)v(x)\psi(x)=g(x)\varphi(x)v(x)=g(x)u(x)\psi(x)$$

и поэтому, после сокращения на не равный нулю (как знаменатель одной из дробей) многочлен $\psi(x)$, получаем:

$$f(x)v(x)=g(x)u(x),$$

откуда, по определению равенства дробей,

$$\frac{f(x)}{g(x)}=\frac{u(x)}{v(x)},$$

что и требовалось доказать.

Объединим теперь в один класс все дроби, равные некоторой данной, и поэтому, в силу транзитивности равенства, равные между собой. Если в одном классе имеется хотя бы одна дробь, не

содержащаяся в другом классе, то, как следует из транзитивности равенства, эти два класса не имеют ни одного общего элемента.

Таким образом, совокупность всех рациональных дробей, написанных при помощи многочленов из кольца $P[x]$, распадается на непересекающиеся классы равных между собой дробей. Мы хотим теперь так определить алгебраические операции в этом множестве классов равных дробей, чтобы оно оказалось полем. Для этого мы будем определять операции над рациональными дробями и каждый раз проверять, что замена слагаемых (или множителей) равными им дробями заменяет сумму (или произведение) также равной дробью. Это позволит говорить о сумме и произведении классов равных дробей.

Предварительно сделаем следующее замечание, которое дальше будет неоднократно применяться: *рациональная дробь превращается в равную дробь, если ее числитель и знаменатель умножаются на один и тот же многочлен, отличный от нуля, или же сокращаются на любой общий множитель.* Действительно,

$$\frac{f(x)}{g(x)}=\frac{f(x)\,h(x)}{g(x)\,h(x)},$$

так как в кольце $P[x]$

$$f(x)\,[g(x)\,h(x)]=g(x)\,[f(x)\,h(x)].$$

Сложение рациональных дробей мы определяем по формуле (2); так как из $g(x)\neq 0$ и $\psi(x)\neq 0$ следует $g(x)\,\psi(x)\neq 0$, то правая часть этой формулы действительно будет рациональной дробью. Если дано, далее, что

$$\frac{f(x)}{g(x)}=\frac{f_0(x)}{g_0(x)},\qquad \frac{\varphi(x)}{\psi(x)}=\frac{\varphi_0(x)}{\psi_0(x)},$$

т. е.

$$f(x)\,g_0(x)=g(x)\,f_0(x),\qquad \varphi(x)\,\psi_0(x)=\psi(x)\,\varphi_0(x), \tag{6}$$

то, умножая обе части первого из равенств (6) на $\psi(x)\,\psi_0(x)$, обе части второго равенства — на $g(x)\,g_0(x)$, а затем складывая эти равенства почленно, мы получим:

$$[f(x)\,\psi(x)+g(x)\,\varphi(x)]\,g_0(x)\,\psi_0(x)=$$
$$=[f_0(x)\,\psi_0(x)+g_0(x)\,\varphi_0(x)]\,g(x)\,\psi(x),$$

что равносильно равенству

$$\frac{f(x)\,\psi(x)+g(x)\,\varphi(x)}{g(x)\,\psi(x)}=\frac{f_0(x)\,\psi_0(x)+g_0(x)\,\varphi_0(x)}{g_0(x)\,\psi_0(x)}.$$

Таким образом, если даны два класса равных между собой дробей, то суммы любой дроби из одного класса с любой дробью из другого класса все между собой равны, т. е. лежат в некотором вполне определенном третьем классе. Этот класс называется *суммой* заданных двух классов.

Коммутативность этого сложения непосредственно вытекает из (2), а ассоциативность доказывается следующим образом:

$$\left[\frac{f(x)}{g(x)}+\frac{\varphi(x)}{\psi(x)}\right]+\frac{u(x)}{v(x)}=\frac{f(x)\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}+\frac{u(x)}{v(x)}=$$
$$=\frac{f(x)\psi(x)v(x)+g(x)\varphi(x)v(x)+g(x)\psi(x)u(x)}{g(x)\psi(x)v(x)}=$$
$$=\frac{f(x)}{g(x)}+\frac{\varphi(x)v(x)+\psi(x)u(x)}{\psi(x)v(x)}=\frac{f(x)}{g(x)}+\left[\frac{\varphi(x)}{\psi(x)}+\frac{u(x)}{v(x)}\right].$$

Из определения равенства дробей без труда следует, что все дроби вида $\frac{0}{g(x)}$, т. е. дроби с равным нулю числителем, равны между собой и что они составляют полный класс равных дробей. Этот класс мы назовем *нулевым* и докажем, что он играет в нашем сложении роль нуля. Действительно, если дана произвольная дробь $\frac{\varphi(x)}{\psi(x)}$, то

$$\frac{0}{g(x)}+\frac{\varphi(x)}{\psi(x)}=\frac{0\cdot\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}=\frac{g(x)\varphi(x)}{g(x)\psi(x)}=\frac{\varphi(x)}{\psi(x)}.$$

Из равенства

$$\frac{f(x)}{g(x)}+\frac{-f(x)}{g(x)}=\frac{0}{g^2(x)},$$

правая часть которого принадлежит к нулевому классу, следует теперь, что класс дробей, равных дроби $\frac{-f(x)}{g(x)}$, будет *противоположным* для класса дробей, равных дроби $\frac{f(x)}{g(x)}$. Отсюда, как мы знаем, следует выполнимость однозначного *вычитания*.

Умножение рациональных дробей мы определим по формуле (3), причем, ввиду $g(x)\psi(x)\neq 0$, правая часть этой формулы действительно будет рациональной дробью. Если, далее,

$$\frac{f(x)}{g(x)}=\frac{f_0(x)}{g_0(x)},\qquad \frac{\varphi(x)}{\psi(x)}=\frac{\varphi_0(x)}{\psi_0(x)},$$

т. е.

$$f(x)g_0(x)=g(x)f_0(x),\qquad \varphi(x)\psi_0(x)=\psi(x)\varphi_0(x),$$

то, перемножая эти последние равенства почленно, мы получим:

$$f(x)g_0(x)\varphi(x)\psi_0(x)=g(x)f_0(x)\psi(x)\varphi_0(x),$$

что равносильно равенству

$$\frac{f(x)\varphi(x)}{g(x)\psi(x)}=\frac{f_0(x)\varphi_0(x)}{g_0(x)\psi_0(x)}.$$

Таким образом, по аналогии с данным выше определением суммы классов, можно говорить о *произведении* классов равных между собой дробей.

Коммутативность и ассоциативность этого умножения непосредственно следуют из (3), а справедливость закона дистрибутивности доказывается следующим образом:

$$\left[\frac{f(x)}{g(x)}+\frac{\varphi(x)}{\psi(x)}\right]\frac{u(x)}{v(x)}=\frac{f(x)\psi(x)+g(x)\varphi(x)}{g(x)\psi(x)}\cdot\frac{u(x)}{v(x)}=$$
$$=\frac{[f(x)\psi(x)+g(x)\varphi(x)]\,u(x)}{g(x)\psi(x)v(x)}=\frac{f(x)\psi(x)u(x)+g(x)\varphi(x)u(x)}{g(x)\psi(x)v(x)}=$$
$$=\frac{f(x)\psi(x)u(x)v(x)+g(x)\varphi(x)u(x)v(x)}{g(x)\psi(x)v^2(x)}=\frac{f(x)u(x)}{g(x)v(x)}+\frac{\varphi(x)u(x)}{\psi(x)v(x)}=$$
$$=\frac{f(x)}{g(x)}\cdot\frac{u(x)}{v(x)}+\frac{\varphi(x)}{\psi(x)}\cdot\frac{u(x)}{v(x)}.$$

Легко видеть, что дроби вида $\frac{f(x)}{f(x)}$, т. е. дроби, числитель которых равен знаменателю, все равны между собой и составляют отдельный класс. Этот класс называется *единичным* и играет в нашем умножении роль единицы:

$$\frac{f(x)}{f(x)}\cdot\frac{\varphi(x)}{\psi(x)}=\frac{f(x)\varphi(x)}{f(x)\psi(x)}=\frac{\varphi(x)}{\psi(x)}.$$

Если, наконец, дробь $\frac{f(x)}{g(x)}$ не принадлежит к нулевому классу, т. е. $f(x)\neq 0$, то существует дробь $\frac{g(x)}{f(x)}$. Так как

$$\frac{f(x)}{g(x)}\cdot\frac{g(x)}{f(x)}=\frac{f(x)g(x)}{g(x)f(x)},$$

а правая часть этого равенства принадлежит к единичному классу, то класс дробей, равных дроби $\frac{g(x)}{f(x)}$, будет *обратным* для класса дробей, равных дроби $\frac{f(x)}{g(x)}$. Отсюда следует выполнимость однозначного *деления*.

Таким образом, *классы равных между собой рациональных дробей с коэффициентами из поля P составляют при нашем определении операций коммутативное поле*. Это поле и будет искомым полем $P(x)$. Мы должны еще, впрочем, доказать, что в построенном нами поле содержится подкольцо, изоморфное кольцу $P[x]$, и что всякий элемент поля представим в виде частного двух элементов из этого подкольца.

Если мы произвольному многочлену $f(x)$ из кольца $P[x]$ поставим в соответствие класс рациональных дробей, равных дроби $\frac{f(x)}{1}$ (среди всех дробей содержатся, понятно, и дроби, знаменатель которых равен единице), то получим взаимно однозначное отображение

кольца $P[x]$ внутрь построенного нами поля. Действительно, из равенства

$$\frac{f(x)}{1}=\frac{\varphi(x)}{1}$$

следовало бы $f(x)\cdot 1=1\cdot\varphi(x)$, т. е. $f(x)=\varphi(x)$. Это отображение будет даже изоморфным, как показывают равенства

$$\frac{f(x)}{1}+\frac{g(x)}{1}=\frac{f(x)\cdot 1+g(x)\cdot 1}{1^2}=\frac{f(x)+g(x)}{1},$$

$$\frac{f(x)}{1}\cdot\frac{g(x)}{1}=\frac{f(x)\cdot g(x)}{1}.$$

Таким образом, *классы дробей, равных дробям вида* $\frac{f(x)}{1}$, *составляют в нашем поле подкольцо, изоморфное кольцу* $P[x]$. Дробь $\frac{f(x)}{1}$ можно поэтому обозначить просто $f(x)$. Так как, наконец, при $g(x)\neq 0$ класс дробей, равных дроби $\frac{1}{g(x)}$, является обратным для класса дробей, равных дроби $\frac{g(x)}{1}$, то из равенства

$$\frac{f(x)}{1}\cdot\frac{1}{g(x)}=\frac{f(x)}{g(x)}$$

следует, что *все элементы нашего поля можно считать* (в смысле операций, определенных в этом поле) *частными многочленов из кольца* $P[x]$.

Мы построили над произвольным полем P поле рациональных дробей $P(x)$. Этим же методом, беря вместо кольца многочленов кольцо целых чисел, можно построить поле рациональных чисел. Объединяя эти два случая и используя такой же метод, можно было бы доказать теорему, что вообще всякое коммутативное кольцо без делителей нуля является подкольцом некоторого поля.

ГЛАВА ОДИННАДЦАТАЯ

МНОГОЧЛЕНЫ ОТ НЕСКОЛЬКИХ НЕИЗВЕСТНЫХ

§ 51. Кольцо многочленов от нескольких неизвестных

Нередко приходится рассматривать многочлены, зависящие не от одного, а от двух, трех, вообще от нескольких неизвестных. Так, в первых главах книги нами уже изучались линейные и квадратичные формы, представляющие собой примеры таких многочленов. Вообще *многочленом* $f(x_1, x_2, \ldots, x_n)$ *от n неизвестных* $x_1, x_2, \ldots, x_n$ *над некоторым полем* P называется сумма конечного числа членов вида $x_1^{k_1}x_2^{k_2} \ldots x_n^{k_n}$, где все $k_i \geqslant 0$, с коэффициентами из поля P; при этом предполагается, понятно, что многочлен $f(x_1, x_2, \ldots, x_n)$ не содержит подобных членов и что рассматриваются лишь члены с отличными от нуля коэффициентами. Два многочлена от n неизвестных, $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$, считаются *равными* (или *тождественно равными*), если равны их коэффициенты при одинаковых членах.

Если дан многочлен $f(x_1, x_2, \ldots, x_n)$ над полем P, то его *степенью по отношению к неизвестному* x_i, $i=1,2, \ldots, n$, называется наивысший показатель, с каким входит x_i в члены этого многочлена. Случайно эта степень может быть равной 0, что означает, что хотя f считается многочленом от n неизвестных $x_1, x_2, \ldots, x_i, \ldots, x_n$, но неизвестное x_i на самом деле в его запись не входит.

С другой стороны, если мы назовем *степенью члена*

$$x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$$

число $k_1+k_2+\ldots+k_n$, т. е. сумму показателей при неизвестных, то *степенью многочлена* $f(x_1, x_2, \ldots, x_n)$ (т. е. степенью по совокупности неизвестных) будет наивысшая из степеней его членов. В частности, многочленами нулевой степени будут, как и в случае одного неизвестного, лишь отличные от нуля элементы из поля P. С другой стороны, как и в случае многочленов от одного неизвестного, нуль будет единственным многочленом от n неизвестных, степень которого не определена. Понятно, что многочлен в общем случае может содержать несколько членов наивысшей степени и поэтому нельзя говорить о старшем (по степени) члене многочлена.

Для многочленов от n неизвестных над полем P следующим образом определяются операции сложения и умножения. *Суммой* многочленов $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ называется многочлен, коэффициенты которого получаются сложением соответственных коэффициентов многочленов f и g; если при этом некоторый член входит лишь в один из многочленов f, g, то коэффициент при нем в другом многочлене считается, понятно, равным нулю. Произведение двух «одночленов» определяется следующим равенством:

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\cdot bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}=(ab)\,x_1^{k_1+l_1}x_2^{k_2+l_2}\ldots x_n^{k_n+l_n},$$

после чего *произведение* многочленов $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ определяется как результат почленного перемножения и последующего приведения подобных членов.

При таком определении операций совокупность многочленов от n неизвестных над полем P превращается в коммутативное кольцо, причем это кольцо не содержит делителей нуля. В самом деле, при $n=1$ наши определения совпадают с теми, которые были даны в § 20 для случая многочленов от одного неизвестного. Пусть уже доказано, что многочлены от $n-1$ неизвестных $x_1, x_2, \ldots, x_{n-1}$ с коэффициентами из поля P составляют кольцо без делителей нуля. Всякий многочлен от n неизвестных $x_1, x_2, \ldots, x_{n-1}, x_n$ можно представить, притом единственным способом, как многочлен от неизвестного x_n с коэффициентами, являющимися многочленами от $x_1, x_2, \ldots, x_{n-1}$; обратно, всякий многочлен от x_n с коэффициентами из кольца многочленов от $x_1, x_2, \ldots, x_{n-1}$ над полем P можно рассматривать, конечно, как многочлен над этим же полем P от всей совокупности неизвестных $x_1, x_2, \ldots, x_{n-1}, x_n$. Без труда проверяется, что полученное нами взаимно однозначное соответствие между многочленами от n неизвестных и многочленами от одного неизвестного над кольцом многочленов от $n-1$ неизвестных является изоморфным по отношению к операциям сложения и умножения. Доказываемое утверждение вытекает теперь из того, что многочлены от одного неизвестного над кольцом многочленов от $n-1$ неизвестных сами составляют кольцо, причем оно как кольцо многочленов от одного неизвестного над кольцом без делителей нуля само не содержит делителей нуля (см. § 47).

Мы доказали, следовательно, существование *кольца многочленов от n неизвестных над полем P;* это кольцо обозначается символом $P[x_1, x_2, \ldots, x_n]$.

Следующие рассмотрения позволяют посмотреть на кольцо многочленов от n неизвестных с несколько иной точки зрения. Пусть поле P содержится в некотором коммутативном кольце L в качестве подкольца. Возьмем в L n элементов $\alpha_1, \alpha_2, \ldots, \alpha_n$ и найдем минимальное подкольцо L' кольца L, содержащее эти элементы и все поле P, т. е. подкольцо, получающееся в результате *присоединения*

к полю P элементов $\alpha_1, \alpha_2, \ldots, \alpha_n$. Подкольцо L' состоит из всех элементов кольца L, которые выражаются через элементы $\alpha_1, \alpha_2, \ldots, \alpha_n$ и элементы поля P при помощи сложения, вычитания и умножения. Легко видеть, что это будут в точности те элементы кольца L, которые можно записать (при помощи операций, имеющих место в L) в виде многочленов от $\alpha_1, \alpha_2, \ldots, \alpha_n$ с коэффициентами из P, причем эти элементы будут как элементы кольца L между собой складываться и умножаться как раз по указанным выше правилам сложения и умножения многочленов от n неизвестных.

Конечно, данный элемент β из подкольца L' будет, вообще говоря, обладать многими различными записями в виде многочлена от $\alpha_1, \alpha_2, \ldots, \alpha_n$ с коэффициентами из поля P. Если для всякого β из L' такая запись однозначна, т. е. если различные многочлены от $\alpha_1, \alpha_2, \ldots, \alpha_n$ будут различными элементами кольца L' (и, следовательно, кольца L), то система элементов $\alpha_1, \alpha_2, \ldots, \alpha_n$ называется *алгебраически независимой* над полем P, в противном случае — *алгебраически зависимой*[1]). Отсюда можно вывести такое заключение:

Если поле P является подкольцом коммутативного кольца L и если система элементов $\alpha_1, \alpha_2, \ldots, \alpha_n$ из L алгебраически независима над P, то подкольцо L' кольца L, порождаемое присоединением к полю P элементов $\alpha_1, \alpha_2, \ldots, \alpha_n$, изоморфно кольцу многочленов $P[x_1, x_2, \ldots, x_n]$.

Из других свойств кольца многочленов от n неизвестных $P[x_1, x_2, \ldots, x_n]$ укажем на следующее: это кольцо можно включить в *поле рациональных дробей $P(x_1, x_2, \ldots, x_n)$ от n неизвестных над полем P*. Всякий элемент этого поля может быть записан в виде $\frac{f}{g}$, где f и g — многочлены из кольца $P[x_1, x_2, \ldots, x_n]$, причем тогда и только тогда $\frac{f}{g} = \frac{\varphi}{\psi}$, если $f\psi = g\varphi$. Сложение и умножение этих рациональных дробей производятся по правилам, которые, как было указано в § 45, справедливы для частных во всяком поле. Доказательство существования поля $P(x_1, x_2, \ldots, x_n)$ проводится так же, как это делалось в § 50 для случая $n = 1$.

Для многочленов от нескольких неизвестных можно построить теорию делимости, обобщающую ту теорию делимости для многочленов от одного неизвестного, которую мы изучали в гл. 5 и 10. Так как, однако, детальное изучение кольца многочленов от нескольких неизвестных не входит в наши задачи, то мы ограничимся только вопросом о разложении многочлена на неприводимые множители.

Введем сначала следующее понятие: если все члены многочлена $f(x_1, x_2, \ldots, x_n)$ имеют одну и ту же степень s, то такой многочлен называется *однородным многочленом* или, короче, *формой* s-й степени; нам

[1]) Соответствующие понятия для случая $n = 1$ были уже введены в § 47: элемент α, алгебраически независимый над полем P в смысле только что данного определения, был назван там трансцендентным над P, в противном случае — алгебраическим над P.

уже известны *линейные* и *квадратичные формы*, можно рассматривать, далее, *кубичные формы*, все члены которых имеют по совокупности неизвестных степень 3, и т. д. Всякий многочлен от n неизвестных однозначно представим в виде суммы нескольких форм от этих неизвестных, притом имеющих разные степени: достаточно объединить вместе все члены, имеющие одну и ту же степень, чтобы получить искомое представление. Так, многочлен четвертой степени $f(x_1, x_2, x_3) = 3x_1x_2^2 - 7x_1^2x_3^2 + x_2 - 5x_1x_2x_3 + x_1^4 - 2x_3 - 6 + x_3^3$ будет суммой формы четвертой степени $x_1^4 - 7x_1^2x_3^2$, кубичной формы $3x_1x_2^2 - 5x_1x_2x_3 + x_3^3$, линейной формы $x_2 - 2x_3$ и свободного члена (формы нулевой степени) -6.

Докажем теперь следующую теорему:

Степень произведения двух отличных от нуля многочленов от n неизвестных равна сумме степеней этих многочленов.

Предположим сначала, что нам даны формы $\varphi(x_1, x_2, \ldots, x_n)$ степени s и $\psi(x_1, x_2, \ldots, x_n)$ степени t. Произведение любого члена формы φ на любой член формы ψ будет, очевидно, иметь степень $s+t$, а потому произведение $\varphi\psi$ будет формой степени $s+t$, так как приведение подобных членов не может сделать все коэффициенты этого произведения равными нулю ввиду отсутствия в кольце $P[x_1, x_2, \ldots, x_n]$ делителей нуля.

Если теперь даны произвольные многочлены $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$ соответственно степеней s и t, то, представляя каждый из них в виде суммы форм разных степеней, мы получим:

$$f(x_1, x_2, \ldots, x_n) = \varphi(x_1, x_2, \ldots, x_n) + \ldots,$$

$$g(x_1, x_2, \ldots, x_n) = \psi(x_1, x_2, \ldots, x_n) + \ldots,$$

где φ и ψ будут соответственно формами степеней s и t, а многоточия заменяют суммы форм меньших степеней. Тогда

$$fg = \varphi\psi + \ldots;$$

форма $\varphi\psi$ имеет, по доказанному, степень $s+t$, а так как все члены, замененные многоточием, имеют меньшую степень, то степень произведения fg будет равна $s+t$. Теорема доказана.

Многочлен φ называется *делителем* многочлена f, а f — *делящимся на* φ, если в кольце $P[x_1, x_2, \ldots, x_n]$ существует такой многочлен ψ, что $f = \varphi\psi$. Легко видеть, что свойства делимости I—IX из § 21 сохраняются и в рассматриваемом сейчас общем случае. Многочлен f степени k, $k \geqslant 1$, называется *приводимым* над полем P, если он разлагается в произведение многочленов из кольца $P[x_1, x_2, \ldots, x_n]$, степени которых меньше k, и *неприводимым* — в противоположном случае.

Всякий многочлен из кольца $P[x_1, x_2, \ldots, x_n]$, имеющий степень, отличную от нуля, разлагается в произведение неприводимых множителей. Это разложение однозначно с точностью до множителей нулевой степени.

Эта теорема обобщает соответствующие результаты из § 48, относящиеся к многочленам от одного неизвестного. Ее первое утверждение доказывается дословным повторением рассуждений из указанного параграфа. Доказательство второго утверждения представляет уже значительные трудности. Прежде чем проводить его, мы заметим, что из второго утверждения этой теоремы вытекает такое следствие: *если произведение двух многочленов f и g из кольца $P[x_1, x_2, \ldots, x_n]$ делится на неприводимый многочлен p, то хотя бы один из этих многочленов делится на p.* Действительно, в противном случае мы получили бы для произведения fg два разложения на неприводимые множители, одно из которых p не содержит, а другое содержит.

Пусть теорема уже доказана для многочленов от n неизвестных, и мы хотим доказать ее для многочлена от $n+1$ неизвестных $x, x_1, x_2, \ldots, x_n$. Запишем этот многочлен в виде $\varphi(x)$; его коэффициенты будут, следовательно, многочленами от $x_1, x_2, \ldots, x_n$. Для этих коэффициентов теорема уже доказана, т. е. каждый из них однозначно разлагается в произведение неприводимых множителей. Назовем многочлен $\varphi(x)$ *примитивным* (точнее, *примитивным над кольцом* $P[x_1, x_2, \ldots, x_n]$), если его коэффициенты не содержат ни одного общего неприводимого множителя, т. е. в совокупности взаимно просты, и докажем следующую лемму Гаусса:

Произведение двух примитивных многочленов само есть примитивный многочлен.

В самом деле, пусть даны примитивные многочлены

$$f(x)=a_0x^k+a_1x^{k-1}+\ldots+a_ix^{k-i}+\ldots+a_k,$$
$$g(x)=b_0x^l+b_1x^{l-1}+\ldots+b_jx^{l-j}+\ldots+b_l$$

с коэффициентами из кольца $P[x_1, x_2, \ldots, x_n]$ и пусть

$$f(x)\,g(x)=c_0x^{k+l}+c_1x^{k+l-1}+\ldots+c_{i+j}x^{k+l-(i+j)}+\ldots+c_{k+l}.$$

Если это произведение не примитивно, то коэффициенты $c_0, c_1, \ldots, c_{k+l}$ будут обладать общим неприводимым множителем $p=p(x_1, x_2, \ldots, x_n)$. Так как все коэффициенты примитивного многочлена $f(x)$ не могут делиться на p, то пусть коэффициент a_i будет первым, на p не делящимся; аналогично через b_j мы обозначим первый коэффициент многочлена $g(x)$, не делящийся на p. Перемножая почленно $f(x)$ и $g(x)$ и собирая члены, содержащие $x^{k+l-(i+j)}$, мы получим:

$$c_{i+j}=a_ib_j+a_{i-1}b_{j+1}+a_{i-2}b_{j+2}+\ldots+a_{i+1}b_{j-1}+a_{i+2}b_{j-2}+\ldots$$

Левая часть этого равенства делится на неприводимый многочлен p. На него заведомо делятся также все слагаемые правой части, кроме первого; действительно, ввиду условий, наложенных на выбор i и j, все коэффициенты $a_{i-1}, a_{i-2}, \ldots$, а также $b_{j-1}, b_{j-2}, \ldots$ делятся на p. Отсюда следует, что произведение a_ib_j также делится на p, а поэтому, как отмечено выше, на p должен делиться хотя бы один из многочленов a_i, b_j, что, однако, не имеет места. Этим заканчивается доказательство леммы в предположении справедливости основной теоремы для многочленов от n неизвестных.

Кольцо $P[x_1, x_2, \ldots, x_n]$ содержится, как мы знаем, в поле рациональных дробей $P(x_1, x_2, \ldots, x_n)$, которое мы обозначим через Q:

$$Q=P(x_1, x_2, \ldots, x_n).$$

Рассмотрим кольцо многочленов $Q[x]$. Если многочлен $\varphi(x)$ принадлежит к этому кольцу, то каждый его коэффициент представим в виде частного многочленов из кольца $P[x_1, x_2, \ldots, x_n]$. Вынося за скобки общий знаменатель этих частных, а затем и общие множители из числителей, можно представить $\varphi(x)$ в виде

$$\varphi(x)=\frac{a}{b}f(x).$$

Здесь a и b являются многочленами из кольца $P[x_1, x_2, \ldots, x_n]$, а $f(x)$ — многочленом от x с коэффициентами из $P[x_1, x_2, \ldots, x_n]$, причем даже примитивным многочленом, так как его коэффициенты уже не имеют общих множителей.

Этим путем всякому многочлену $\varphi(x)$ из кольца $Q[x]$ поставлен в соответствие примитивный многочлен $f(x)$. *Для данного $\varphi(x)$ многочлен $f(x)$*

определен однозначно с точностью до отличного от нуля множителя из поля P. Действительно, пусть

$$\varphi(x)=\frac{a}{b}f(x)=\frac{c}{d}g(x),$$

где $g(x)$ — снова примитивный многочлен. Тогда

$$adf(x)=bcg(x).$$

Таким образом, ad и bc получены вынесением всех общих множителей из коэффициентов одного и того же многочлена над кольцом $P[x_1, x_2, \ldots, x_n]$. Отсюда вытекает, ввиду справедливости в этом кольце (по предположению индукции) теоремы об однозначности разложения, что ad и bc могут отличаться друг от друга лишь множителем нулевой степени. Таким же множителем, следовательно, отличаются друг от друг примитивные многочлены $f(x)$ и $g(x)$.

Произведению двух многочленов из кольца $Q[x]$ соответствует произведение соответствующих им примитивных многочленов. В самом деле, если

$$\varphi(x)=\frac{a}{b}f(x), \quad \psi(x)=\frac{c}{d}g(x),$$

где $f(x)$ и $g(x)$ — примитивные многочлены, то

$$\varphi(x)\psi(x)=\frac{ac}{bd}f(x)g(x).$$

Но, как доказано выше, произведение $f(x)g(x)$ является примитивным многочленом.

Отметим, далее, что *если многочлен $\varphi(x)$ из кольца $Q[x]$ неприводим над полем Q, то соответствующий ему примитивный многочлен $f(x)$, рассматриваемый как многочлен от $x, x_1, x_2, \ldots, x_n$, также будет неприводимым, и обратно*. В самом деле, если многочлен f приводим, $f=f_1f_2$, то оба множителя должны содержать неизвестное x, так как иначе многочлен f не был бы примитивным. Отсюда вытекает разложение многочлена $\varphi(x)$ над полем Q:

$$\varphi(x)=\frac{a}{b}f(x)=\left(\frac{a}{b}f_1\right)f_2.$$

Обратно, если многочлен $\varphi(x)$ приводим над Q, $\varphi(x)=\varphi_1(x)\varphi_2(x)$, то примитивные многочлены $f_1(x)$ и $f_2(x)$, соответствующие многочленам $\varphi_1(x)$ и $\varphi_2(x)$, будут оба содержать x, но их произведение, как доказано выше, равно $f(x)$ (с точностью до множителя из поля P).

Возьмем теперь примитивный многочлен f и разложим его на неприводимые множители, $f=f_1 \cdot f_2 \ldots f_k$. Все эти множители не только должны содержать неизвестное x, но даже будут примитивными многочленами, так как иначе многочлен f не был бы примитивным. *Это разложение примитивного многочлена f будет однозначным с точностью до множителей из поля P*. В самом деле, ввиду предшествующей леммы, можно смотреть на это разложение как на разложение $f(x)$ на неприводимые множители над полем Q, но для многочленов от одного неизвестного над некоторым полем однозначность разложения нам уже известна; эта однозначность имеет место с точностью до множителей из Q; однако в нашем случае, благодаря примитивности всех множителей f_i, она будет с точностью до множителей из P.

После этих лемм, справедливость которых нами доказана, исходя из индуктивного предположения, доказательство нашей основной теоремы проходит без всяких затруднений. В самом деле, всякий неприводимый многочлен из кольца $P[x, x_1, x_2, \ldots, x_n]$ будет или неприводимым многочленом

из кольца $P[x_1, x_2, \ldots, x_n]$, или же неприводимым примитивным многочленом. Отсюда следует, что если нам дано некоторое разложение многочлена $\varphi(x, x_1, x_2, \ldots, x_n)$ на неприводимые множители, то, объединяя множители, мы представим φ в виде

$$\varphi(x, x_1, x_2, \ldots, x_n) = a(x_1, x_2, \ldots, x_n)\, f(x, x_1, x_2, \ldots, x_n),$$

где a от x не зависит, а f является примитивным многочленом. Мы знаем, однако, что это разложение для φ однозначно с точностью до множителей из P. Так как, с другой стороны, однозначность разложения на неприводимые множители для многочлена a от n неизвестных имеет место по предположению индукции, а для примитивного многочлена f доказана в предшествующей лемме, то наша теорема для случая $n+1$ неизвестных также полностью доказана.

Из доказанных выше лемм вытекает еще одно интересное следствие: *если многочлен $\varphi(x)$ с коэффициентами из $P[x_1, x_2, \ldots, x_n]$ приводим над полем $Q = P(x_1, x_2, \ldots, x_n)$, то он может быть разложен на множители, зависящие от x и имеющие коэффициентами многочлены из кольца $P[x_1, x_2, \ldots, x_n]$.* Действительно, если многочлену $\varphi(x)$ соответствует примитивный многочлен $f(x)$, т. е. $\varphi(x) = af(x)$, то, как мы знаем, из разложимости $\varphi(x)$ следует разложимость $f(x)$; последнее приводит, однако, к разложению $\varphi(x)$ над кольцом $P[x_1, x_2, \ldots, x_n]$.

В отличие от случая многочленов от одного неизвестного, которые, как мы знаем из § 49, могут быть разложены на линейные множители над соответственно подобранным расширением рассматриваемого основного поля, *над любым полем P существуют абсолютно неприводимые многочлены произвольной степени от нескольких* (двух или более) *неизвестных*, т. е. многочлены, которые остаются неприводимыми при любом расширении этого поля.

Таков, например, многочлен

$$f(x, y) = \varphi(x) + y,$$

где $\varphi(x)$ — произвольный многочлен от одного неизвестного над полем P. Действительно, если бы в некотором расширении $\bar{P}$ поля P существовало разложение

$$f(x, y) = g(x, y)\, h(x, y),$$

то, записывая g и h по степеням y, мы получили бы, что, например,

$$g(x, y) = a_0(x)\, y + a_1(x), \quad h(x, y) = b_0(x),$$

т. е. h не зависит от y, а затем, ввиду $a_0(x)\, b_0(x) = 1$, что $b_0(x)$ имеет степень 0, т. е. h не зависит и от x.

Лексикографическое расположение членов многочлена. Для многочленов от одного неизвестного мы имеем два естественных способа расположения членов — по убывающим и по возрастающим степеням неизвестного. В случае многочленов от нескольких неизвестных такие способы уже отсутствуют: если дан многочлен пятой степени от трех неизвестных

$$f(x_1, x_2, x_3) = x_1x_2^2x_3^2 + x_1^4x_3 + x_2^3x_3^2 + x_1^2x_2x_3^2,$$

то его можно было бы записать и в виде

$$f(x_1, x_2, x_3) = x_1^4x_3 + x_1^2x_2x_3^2 + x_1x_2^2x_3^2 + x_2^3x_3^2,$$

и нет оснований одну из этих записей предпочесть другой. Существует, однако, способ вполне определенного расположения членов многочлена от нескольких неизвестных, зависящий, впрочем, от выбора нумерации неизвестных; для многочленов от одного неизвестного он приводит к расположению членов по убывающим степеням неизвестного. Этот способ, называемый *лексикографическим*, подсказан обычным приемом расположения слов в словарях («лексиконах»): считая буквы упорядоченными так, как это принято в алфавите, мы определяем взаимное положение двух данных слов в словаре по их первым буквам, если же эти буквы совпадают, то по вторым буквам и т. д.

Пусть дан многочлен $f(x_1, x_2, \ldots, x_n)$ из кольца $P[x_1, x_2, \ldots, x_n]$ и в нем два различных члена:

$$x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, \tag{1}$$

$$x_1^{l_1} x_2^{l_2} \ldots x_n^{l_n}, \tag{2}$$

коэффициенты которых являются некоторыми отличными от нуля элементами из P. Так как члены (1) и (2) различны, то хотя бы одна из разностей показателей при неизвестных

$$k_i - l_i, \quad i = 1, 2, \ldots, n,$$

отлична от нуля. Член (1) будет считаться *выше* члена (2) (а член (2) — *ниже* члена (1)), если первая из этих разностей, отличная от нуля, положительна, т. е. если существует такое i, $1 \leqslant i \leqslant n$, что

$$k_1 = l_1, \quad k_2 = l_2, \ldots, k_{i-1} = l_{i-1}, \text{ но } k_i > l_i.$$

Иными словами, член (1) будет выше члена (2), если показатель при x_1 в (1) больше, чем в (2), или если эти показатели равны, но показатель при x_2 в (1) больше, чем в (2), и т. д. Легко видеть, что из того, что член (1) выше члена (2), не следует, что степень первого по совокупности неизвестных больше степени второго: из членов

$$x_1^3 x_2 x_3, \quad x_1 x_2^5 x_3^2$$

первый выше, хотя имеет меньшую степень.

Очевидно, что из любых двух различных членов многочлена $f(x_1, x_2, \ldots, x_n)$ один будет выше другого. Легко проверить также, что если член (1) выше члена (2), а член (2), в свою очередь, выше члена

$$x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n}, \tag{3}$$

т. е. существует такое j, $1 \leqslant j \leqslant n$, что

$$l_1 = m_1, \quad l_2 = m_2, \ldots, l_{j-1} = m_{j-1}, \quad \text{но } l_j > m_j,$$

то, независимо от того, будет ли i больше, равно или меньше j, член (1) будет выше члена (3). Таким образом, ставя раньше тот

из двух членов, который выше, мы получим вполне определенное упорядочение членов многочлена $f(x_1, x_2, \ldots, x_n)$, которое и называется лексикографическим.

Так, многочлен

$$f(x_1, x_2, x_3, x_4)=x_1^4+3x_1^2x_2^3x_3-x_1^2x_2^3x_4^2+5x_1x_3x_4^2+2x_2+x_3^3x_4-4$$

расположен лексикографически.

При лексикографической записи многочлена $f(x_1, x_2, \ldots, x_n)$ один из его членов будет стоять на первом месте, т. е. будет выше всех других членов. Этот член называется *высшим членом многочлена*; в предшествующем примере высшим членом будет член x_1^4. Относительно высших членов мы докажем лемму, которая будет использована при доказательстве основной теоремы следующего параграфа:

Высший член произведения двух многочленов от n неизвестных равен произведению высших членов сомножителей.

В самом деле, пусть перемножаются многочлены $f(x_1, x_2, \ldots, x_n)$ и $g(x_1, x_2, \ldots, x_n)$. Если

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \tag{4}$$

будет высший член многочлена $f(x_1, x_2, \ldots, x_n)$, а

$$a'x_1^{s_1}x_2^{s_2}\ldots x_n^{s_n} \tag{5}$$

— любой другой член этого многочлена, то существует такое i, $1\leqslant i\leqslant n$, что

$$k_1=s_1, \ldots, k_{i-1}=s_{i-1}, \quad k_i>s_i.$$

Если, с другой стороны,

$$bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}, \tag{6}$$

$$b'x_1^{t_1}x_2^{t_2}\ldots x_n^{t_n} \tag{7}$$

будут высший и любой другой члены многочлена $g(x_1, x_2, \ldots, x_n)$, то существует такое j, $1\leqslant j\leqslant n$, что

$$l_1=t_1, \ldots, l_{j-1}=t_{j-1}, \quad l_j>t_j.$$

Перемножая члены (4) и (6), а также члены (5) и (7), мы получаем:

$$abx_1^{k_1+l_1}x_2^{k_2+l_2}\ldots x_n^{k_n+l_n}, \tag{8}$$

$$a'b'x_1^{s_1+t_1}x_2^{s_2+t_2}\ldots x_n^{s_n+t_n}. \tag{9}$$

Легко видеть, однако, что член (8) выше члена (9); если, например, $i\leqslant j$, то

$$k_1+l_1=s_1+t_1, \ldots, k_{i-1}+l_{i-1}=s_{i-1}+t_{i-1}, \quad \text{но} \quad k_i+l_i>s_i+t_i,$$

так как $k_i>s_i$, $l_i\geqslant t_i$. Так же проверяется, что член (8) будет выше произведения членов (4) и (7), а также выше произведения

членов (5) и (6). Таким образом, член (8) — произведение высших членов многочленов f и g — будет выше всех других членов, получающихся в результате почленного перемножения многочленов f и g, а потому этот член не уничтожается при приведении подобных членов, т. е. остается высшим членом в произведении fg.

§ 52. Симметрические многочлены

Среди многочленов от нескольких неизвестных выделяются те, которые не меняются ни при какой перестановке неизвестных. В такие многочлены все неизвестные входят, следовательно, вполне симметричным образом, и поэтому эти многочлены называются *симметрическими многочленами* (или *симметрическими функциями*). Простейшими примерами будут: сумма всех неизвестных $x_1+x_2+\ldots+x_n$, сумма квадратов неизвестных $x_1^2+x_2^2+\ldots$ $\ldots+x_n^2$, произведение неизвестных $x_1x_2\ldots x_n$ и т. д. Ввиду представимости всякой подстановки n символов в виде произведения транспозиций (см. § 3), при доказательстве симметричности некоторого многочлена достаточно проверить, что он не меняется ни при какой транспозиции двух неизвестных.

Мы будем рассматривать дальше симметрические многочлены от n неизвестных с коэффициентами из некоторого поля P. Легко видеть, что *сумма, разность и произведение двух симметрических многочленов сами будут симметрическими*, т. е. симметрические многочлены составляют подкольцо в кольце $P[x_1, x_2, \ldots, x_n]$ всех многочленов от n неизвестных над полем P, называемое ***кольцом симметрических многочленов от n неизвестных над полем P***. К этому кольцу принадлежат все элементы из P (т. е. все многочлены нулевой степени, а также нуль), так как они заведомо не меняются ни при какой перестановке неизвестных. Всякий другой симметрический многочлен непременно содержит все n неизвестных и даже имеет по ним одну и ту же степень: если симметрический многочлен $f(x_1, x_2, \ldots, x_n)$ обладает членом, в который неизвестное x_i входит с показателем k, то обладает и членом, получающимся из него транспозицией неизвестных x_i и x_j, т. е. содержащим неизвестное x_j в той же степени k.

Следующие n симметрических многочленов от n неизвестных называются *элементарными симметрическими многочленами*:

$$\left.\begin{aligned}
\sigma_1 &= x_1+x_2+\ldots+x_n,\\
\sigma_2 &= x_1x_2+x_1x_3+\ldots+x_{n-1}x_n,\\
\sigma_3 &= x_1x_2x_3+x_1x_2x_4+\ldots+x_{n-2}x_{n-1}x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\sigma_{n-1} &= x_1x_2\ldots x_{n-1}+x_1x_2\ldots x_{n-2}x_n+\ldots+x_2x_3\ldots x_n,\\
\sigma_n &= x_1x_2\ldots x_n.
\end{aligned}\right\} \qquad (1)$$

Эти многочлены, симметричность которых очевидна, играют в теории симметрических многочленов очень большую роль. Они подсказаны формулами Вьета (см. § 24), и поэтому можно сказать, что *коэффициенты многочлена от одного неизвестного, имеющего старшим коэффициентом единицу, будут, с точностью до знака, элементарными симметрическими многочленами от его корней.* Эта связь элементарных симметрических многочленов с формулами Вьета будет весьма существенна для тех применений симметрических многочленов к теории многочленов от одного неизвестного, ради которых мы сейчас их изучаем.

Так как симметрические многочлены от n неизвестных $x_1, x_2, \ldots, x_n$ над полем P составляют кольцо, то очевидны следующие утверждения: симметрическим многочленом будет всякая целая положительная степень любого из элементарных симметрических многочленов, а также произведение таких степеней, притом взятое с любым коэффициентом из P, и, наконец, всякая сумма указанных произведений. Иными словами, *всякий многочлен от элементарных симметрических многочленов $\sigma_1, \sigma_2, \ldots, \sigma_n$ с коэффициентами из P, рассматриваемый как многочлен от неизвестных $x_1, x_2, \ldots, x_n$, будет симметрическим.* Так, положим $n=3$ и возьмем многочлен $\sigma_1\sigma_2+2\sigma_3$. Заменяя σ_1, σ_2 и σ_3 их выражениями, мы получим:

$$\sigma_1\sigma_2+2\sigma_3=x_1^2x_2+x_1^2x_3+x_1x_2^2+x_2^2x_3+x_1x_3^2+x_2x_3^2+5x_1x_2x_3;$$

справа стоит, очевидно, симметрический многочлен от x_1, x_2, x_3.

Обращением этого результата является следующая **основная теорема о симметрических многочленах**:

Всякий симметрический многочлен от неизвестных $x_1, x_2, \ldots, x_n$ над полем P является многочленом от элементарных симметрических многочленов $\sigma_1, \sigma_2, \ldots, \sigma_n$ с коэффициентами, принадлежащими к полю P.

Пусть, в самом деле, дан симметрический многочлен

$$f(x_1, x_2, \ldots, x_n)$$

и пусть в его лексикографической записи высшим будет член

$$a_0x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}. \tag{2}$$

Показатели при неизвестных в этом члене должны удовлетворять неравенствам

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_n. \tag{3}$$

Действительно, пусть при некотором i будет $k_i < k_{i+1}$. Многочлен $f(x_1, x_2, \ldots, x_n)$, будучи симметрическим, должен содержать, однако, член

$$a_0x_1^{k_1}x_2^{k_2}\ldots x_i^{k_{i+1}}x_{i+1}^{k_i}\ldots x_n^{k_n}, \tag{4}$$

получающийся из члена (2) транспозицией неизвестных x_i и x_{i+1}. Это приводит нас к противоречию, так как член (4) в смысле лексикографического расположения выше члена (2): показатели при x_1, $x_2, \ldots, x_{i-1}$ в обоих членах совпадают, но показатель при x_i в члене (4) больше, чем в члене (2).

Возьмем теперь следующее произведение элементарных симметрических многочленов (ввиду неравенств (3) все показатели будут неотрицательными):

$$\varphi_1 = a_0\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\ldots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}. \tag{5}$$

Это будет симметрический многочлен от неизвестных $x_1, x_2, \ldots, x_n$, причем его высший член равен члену (2). Действительно, высшие члены многочленов $\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n$ равны соответственно $x_1, x_1x_2, x_1x_2x_3, \ldots, x_1x_2\ldots x_n$, а так как в конце предыдущего параграфа доказано, что высший член произведения равен произведению высших членов сомножителей, то высшим членом многочлена φ_1 будет $a_0x_1^{k_1-k_2}(x_1x_2)^{k_2-k_3}(x_1x_2x_3)^{k_3-k_4}\ldots$

$$\ldots(x_1x_2\ldots x_{n-1})^{k_{n-1}-k_n}(x_1x_2\ldots x_n)^{k_n} = a_0x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}.$$

Отсюда следует, что при вычитании φ_1 из f высшие члены этих многочленов взаимно уничтожатся, т. е. высший член симметрического многочлена $f-\varphi_1=f_1$ будет ниже члена (2), высшего в многочлене f. Повторяя для многочлена f_1, коэффициенты которого принадлежат, очевидно, к полю P, этот же прием, мы придем к равенству

$$f_1 = \varphi_2 + f_2,$$

где φ_2 есть произведение степеней элементарных симметрических многочленов с некоторым коэффициентом из поля P, а f_2 — симметрический многочлен, высший член которого ниже, чем высший член в f_1. Отсюда вытекает равенство

$$f = \varphi_1 + \varphi_2 + f_2.$$

Продолжая этот процесс, мы для некоторого s получим $f_s = 0$ и поэтому придем к выражению для f в виде многочлена от $\sigma_1, \sigma_2, \ldots, \sigma_n$ с коэффициентами из P:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{s}\varphi_i = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n).$$

В самом деле, если бы этот процесс был бесконечным[1]), то мы получили бы бесконечную последовательность симметрических многочленов

$$f_1, f_2, \ldots, f_s, \ldots, \tag{6}$$

[1]) Следует учесть, что многочлен φ_s содержит, вообще говоря, и такие члены, каких нет в многочлене f_{s-1}, и поэтому переход от f_{s-1} к $f_s = f_{s-1} - \varphi_s$ связан не только с уничтожением некоторых членов из f_{s-1}, но и с появлением новых членов. Здесь $s = 1, 2, \ldots$

причем высший член каждого из них был бы ниже, чем высшие члены предшествующих многочленов, и тем более ниже, чем (2). Однако, если

$$bx_1^{l_1}x_2^{l_2}\ldots x_n^{l_n} \qquad (7)$$

есть высший член многочлена f_s, то из симметричности этого многочлена следуют неравенства

$$l_1 \geqslant l_2 \geqslant \ldots \geqslant l_n, \qquad (8)$$

подобные неравенствам (3). С другой стороны, так как член (2) выше члена (7), то

$$k_1 \geqslant l_1. \qquad (9)$$

Легко видеть, однако, что системы целых неотрицательных чисел $l_1, l_2, \ldots, l_n$ удовлетворяющих неравенствам (8) и (9), можно выбрать лишь конечным числом способов. Действительно, если даже отказаться от требования (8) и лишь предполагать, что все l_i, $i = 1, 2, \ldots, n$, не больше k_1, то все равно выбор чисел l_i будет возможен лишь $(k_1+1)^n$ способами. Отсюда следует, что последовательность многочленов (6) со строго понижающимися высшими членами не может быть бесконечной.

Доказательство теоремы закончено.

Отмеченная выше связь элементарных симметрических многочленов с формулами Вьета позволяет вывести такое важное следствие из основной теоремы о симметрических многочленах:

Пусть $f(x)$ есть многочлен от одного неизвестного над полем P, имеющий старшим коэффициентом единицу. Тогда всякий симметрический многочлен (с коэффициентами из P) от корней многочлена $f(x)$, принадлежащих к некоторому полю разложения многочлена $f(x)$ над P, будет многочленом (с коэффициентами из P) от коэффициентов многочлена $f(x)$ и поэтому будет элементом поля P.

Изложенное выше доказательство основной теоремы дает заодно и метод для практического разыскания выражений симметрических многочленов через элементарные. Предварительно введем следующее обозначение: если

$$ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n} \qquad (10)$$

есть некоторое произведение степеней неизвестных $x_1, x_2, \ldots, x_n$ (причем среди показателей могут быть и равные нулю), то через

$$S\left(ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}\right) \qquad (11)$$

будет обозначаться сумма всех членов, получающихся из (10) при всевозможных перестановках неизвестных. Очевидно, что это будет симметрический многочлен, притом однородный, и что всякий симметрический многочлен от n неизвестных, содержащий член (10), будет содержать и все остальные члены многочлена (11). Например, $S(x_1) = \sigma_1$, $S(x_1x_2) = \sigma_2$, $S\left(x_1^2\right)$ есть сумма квадратов всех неизвестных и т. д.

Пример. Симметрический многочлен $f=S(x_1^2x_2)$ от n неизвестных выразить через элементарные симметрические многочлены.

Здесь высший член $x_1^2x_2$ и поэтому $\varphi_1=\sigma_1^{2-1}\sigma_2=\sigma_1\sigma_2$, т. е.

$$\varphi_1=(x_1+x_2+\ldots+x_n)(x_1x_2+x_1x_3+\ldots+x_{n-1}x_n)=$$
$$=S(x_1^2x_2)+3S(x_1x_2x_3),$$

откуда

$$f_1=f-\varphi_1=-3S(x_1x_2x_3)=-3\sigma_3.$$

Поэтому $f=\varphi_1+f_1=\sigma_1\sigma_2-3\sigma_3$.

В более сложных примерах целесообразнее предварительно установить, какие члены могут войти в выражение данного многочлена через элементарные, а затем найти коэффициенты этих членов методом неопределенных коэффициентов.

Примеры. 1. Найти выражение для симметрического многочлена $f=S(x_1^2x_2^2)$.

Мы знаем (см. доказательство основной теоремы), что члены искомого многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ определяются через высшие члены симметрических многочленов $f_1, f_2 \ldots$, причем эти высшие члены ниже высшего члена данного многочлена f, т. е. ниже $x_1^2x_2^2$. Найдем все произведения $x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$, удовлетворяющие следующим условиям: 1) они ниже члена $x_1^2x_2^2$, 2) они могут служить высшими членами симметрических многочленов, т. е. удовлетворяют неравенствам $l_1\geqslant l_2\geqslant\ldots\geqslant l_n$, 3) по совокупности неизвестных они имеют степень 4 (так как все многочлены $f_1, f_2, \ldots$ имеют, как мы знаем, ту же степень, что и однородный многочлен f). Выписывая лишь соответствующие комбинации показателей и указывая рядом те произведения степеней σ, которые ими определяются, мы получаем следующую таблицу:

$$22000\ldots\sigma_1^{2-2}\sigma_2^{2-0}=\sigma_2^2,$$
$$21100\ldots\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-0}=\sigma_1\sigma_3,$$
$$11110\ldots\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-0}=\sigma_4.$$

Таким образом, многочлен f имеет вид

$$f=\sigma_2^2+A\sigma_1\sigma_3+B\sigma_4.$$

Коэффициент при σ_2 мы положили равным единице, так как этот член определен высшим членом многочлена f и имеет, как мы знаем из доказательства основной теоремы, такой же коэффициент. Коэффициенты A и B мы найдем следующим образом.

Положим $x_1=x_2=x_3=1$, $x_4=\ldots=x_n=0$. Легко видеть, что при этих значениях неизвестных многочлен f получает значение 3, а многочлены σ_1, σ_2, σ_3 и σ_4 — соответственно значения 3, 3, 1 и 0. Поэтому

$$3=9+A\cdot 3\cdot 1+B\cdot 0,$$

откуда $A=-2$. Положим теперь $x_1=x_2=x_3=x_4=1$, $x_5=\ldots=x_n=0$. Значения многочленов f, σ_1, σ_2, σ_3 и σ_4 будут равны соответственно 6, 4, 6, 4, 1. Поэтому

$$6=36-2\cdot 4\cdot 4+B\cdot 1,$$

откуда $B=2$. Таким образом, искомое выражение для f будет

$$f=\sigma_2^2-2\sigma_1\sigma_3+2\sigma_4.$$

2. Найти сумму кубов корней многочлена

$$f(x) = x^4 + x^3 + 2x^2 + x + 1.$$

Для решения этой задачи найдем выражение через элементарные симметрические многочлены для симметрического многочлена $S(x_1^3)$. Применяя тот же метод, как и в предыдущем примере, мы получим таблицу

$$\begin{aligned} &3000\ldots\sigma_1^3, \\ &2100\ldots\sigma_1\sigma_2, \\ &1110\ldots\sigma_3, \end{aligned}$$

а поэтому

$$S(x_1^3) = \sigma_1^3 + A\sigma_1\sigma_2 + B\sigma_3.$$

Полагая сперва $x_1 = x_2 = 1$, $x_3 = \ldots = x_n = 0$, а затем $x_1 = x_2 = x_3 = 1$, $x_4 = \ldots = x_n = 0$, мы получим $A = -3$, $B = 3$, т. е.

$$S(x_1^3) = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \qquad (12)$$

Для нахождения суммы кубов корней данного нам многочлена $f(x)$ нужно, ввиду формул Вьета, в найденном выше выражении заменить σ_1 через коэффициент при x^3 с обратным знаком, т. е. через -1, заменить σ_2 через коэффициент при x^2, т. е. через 2, и, наконец, заменить σ_3 через коэффициент при x с обратным знаком, т. е. через -1. Таким образом, интересующая нас сумма кубов корней равна

$$(-1)^3 - 3\cdot(-1)\cdot 2 + 3\cdot(-1) = 2.$$

Читатель может проверить этот результат, если учтет, что $f(x)$ имеет корнями числа i, $-i$, $-\frac{1}{2} + i\frac{\sqrt{3}}{2}$ и $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$. Очевидно также, что формула (12) не зависит от заданного многочлена $f(x)$ и позволяет находить сумму кубов корней любого многочлена.

Метод для выражения симметрического многочлена f через элементарные, полученный при доказательстве основной теоремы, приводит к вполне определенному многочлену от $\sigma_1, \sigma_2, \ldots, \sigma_n$. Оказывается, что никаким способом нельзя получить для f иного выражения через $\sigma_1, \sigma_2, \ldots, \sigma_n$. Это показывает следующая теорема единственности:

Всякий симметрический многочлен обладает лишь единственным выражением в виде многочлена от элементарных симметрических многочленов.

Докажем эту теорему. Если бы симметрический многочлен $f(x_1, x_2, \ldots, x_n)$ на полем P обладал двумя различными выражениями через $\sigma_1, \sigma_2, \ldots, \sigma_n$:

$$f(x_1, x_2, \ldots, x_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \psi(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

то разность

$$\chi(\sigma_1, \sigma_2, \ldots, \sigma_n) = \varphi(\sigma_1, \sigma_2, \ldots, \sigma_n) - \psi(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

была бы отличным от нуля многочленом от $\sigma_1, \sigma_2 \ldots, \sigma_n$, т. е. не все его коэффициенты были бы равны нулю, в то время как замена в этом многочлене $\sigma_1, \sigma_2, \ldots, \sigma_n$ их выражениями через $x_1, x_2, \ldots \ldots, x_n$ приводила бы к нулю кольца $P[x_1, x_2, \ldots, x_n]$. Нам остается поэтому доказать, что если многочлен $\chi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ отличен от нуля, т. е. обладает хотя бы одним отличным от нуля коэффициентом, то и многочлен $g(x_1, x_2, \ldots, x_n)$, полученный из χ заменой $\sigma_1, \sigma_2, \ldots, \sigma_n$ их выражениями через $x_1, x_2, \ldots, x_n$:

$$\chi(\sigma_1, \sigma_2, \ldots, \sigma_n) = g(x_1, x_2, \ldots, x_n), \tag{13}$$

также отличен от нуля.

Если $a\sigma_1^{k_1}\sigma_2^{k_2} \ldots \sigma_n^{k_n}$ — один из членов многочлена χ, причем $a \neq 0$, то после замены всех σ их выражениями (1) мы получим многочлен от $x_1, x_2, \ldots, x_n$, высшим членом которого (в смысле лексикографического расположения) будет, как мы уже знаем из доказательства основной теоремы, член

$$ax_1^{k_1}(x_1x_2)^{k_2} \ldots (x_1x_2 \ldots x_n)^{k_n} = ax_1^{l_1}x_2^{l_2} \ldots x_n^{l_n},$$

где

$$\begin{aligned} l_1 &= k_1 + k_2 + \ldots + k_n, \\ l_2 &= \quad\;\;\; k_2 + \ldots + k_n, \\ &\ldots\ldots\ldots\ldots \\ l_n &= \qquad\qquad\quad\;\; k_n. \end{aligned}$$

Отсюда

$$k_i = l_i - l_{i+1}, \quad k_n = l_n, \qquad i = 1, 2, \ldots, n-1,$$

т. е. по показателям $l_1, l_2, \ldots, l_n$ можно восстановить показатели $k_1, k_2, \ldots, k_n$ исходного члена многочлена χ. Таким образом, различные члены многочлена χ, рассматриваемые как многочлены от $x_1, x_2, \ldots, x_n$, обладают *различными* высшими членами.

Рассмотрим теперь все члены многочлена χ; для каждого из них найдем высший член его представления в виде многочлена от $x_1, x_2, \ldots, x_n$ и отберем тот из этих высших членов, который будет *наивысшим* в смысле лексикографического расположения. Как сказано выше, этот член не имеет подобных среди высших членов, получающихся из других членов многочлена χ, а так как он, по условию, выше каждого из этих *высших* членов, то тем более он выше других членов, получающихся при замене в членах многочлена χ элементов $\sigma_1, \sigma_2, \ldots, \sigma_n$ их выражениями (1). Мы нашли, следовательно, такой член, который при переходе от $\chi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ к $g(x_1, x_2, \ldots, x_n)$ появляется (с отличным от нуля коэффициентом) только один раз и поэтому ни с чем не может сократиться. Отсюда следует, что не все коэффициенты многочлена $g(x_1, x_2, \ldots, x_n)$ равны нулю, т. е. этот многочлен не является нулем кольца $P[x_1, x_2, \ldots, x_n]$, что и требовалось доказать.

Доказанную теорему можно также, очевидно, сформулировать следующим образом:

Система элементарных симметрических многочленов $\sigma_1, \sigma_2, \ldots, \sigma_n$, рассматриваемых как элементы кольца многочленов $P[x_1, x_2, \ldots, x_n]$, алгебраически независима над полем P.

§ 53*. Дополнительные замечания о симметрических многочленах

Замечания к основной теореме. Доказательство основной теоремы о симметрических многочленах, проведенное в предшествующем параграфе, позволяет сделать несколько существенных добавлений к формулировке теоремы, которыми мы ниже воспользуемся. Прежде всего, коэффициенты того многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$, который найден нами в качестве выражения для симметрического многочлена $f(x_1, x_2, \ldots, x_n)$ через элементарные симметрические многочлены, не только принадлежат к полю P, но даже *выражаются через коэффициенты многочлена f при помощи сложения и вычитания, т. е. принадлежат к кольцу L, порождаемому коэффициентами многочлена f внутри поля P.*

В самом деле, все коэффициенты многочлена φ_1 (см. формулу (5) предшествующего параграфа) относительно неизвестных $x_1, x_2, \ldots, x_n$ суть, как легко видеть, целые кратные от коэффициента a_0 при высшем члене многочлена f и поэтому принадлежат к кольцу L. Пусть уже доказано, что к L принадлежат все коэффициенты (относительно $x_1, x_2, \ldots, x_n$) многочленов $\varphi_1, \varphi_2, \ldots, \varphi_l$. Тогда коэффициенты многочлена $f_l = f - \varphi_1 - \varphi_2 - \ldots - \varphi_l$ также будут принадлежать к L, а поэтому в L лежат и все коэффициенты многочлена φ_{l+1} относительно $x_1, x_2, \ldots, x_n$.

С другой стороны, *степень многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ по совокупности $\sigma_1, \sigma_2, \ldots, \sigma_n$ равна степени, которую имеет многочлен $f(x_1, x_2, \ldots, x_n)$ по каждому из неизвестных x_i.* В самом деле, так как (2) из предшествующего параграфа есть высший член многочлена f, то k_1 будет степенью f относительно неизвестного x_1, а поэтому, ввиду симметричности, и относительно любого другого из неизвестных x_i. Однако степень φ_1 по совокупности σ равна, по (5) из предшествующего параграфа, числу

$$(k_1 - k_2) + (k_2 - k_3) + \ldots + (k_{n-1} - k_n) + k_n = k_1.$$

Далее, так как старший член многочлена f_1 ниже старшего члена многочлена f, то степень f_1 по каждому из x_i будет не выше чем степень f по каждому из этих неизвестных. Однако многочлен φ_2 играет для f_1 такую же роль, как φ_1 для f, поэтому степень φ_2 по совокупности σ равна степени f_1 по каждому из x_i, т. е. она не больше чем k_1 и т. д. Таким образом, и степень $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ не выше чем k_1. Поскольку же никакое φ_i с $i > 1$ не может содер-

жать все $\sigma_1, \sigma_2, \ldots, \sigma_n$ в тех же степенях, что и φ_1, то степень $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ в точности равна k_1. Тем самым наше утверждение доказано.

Наконец, пусть $a\sigma_1^{l_1}\sigma_2^{l_2}\ldots\sigma_n^{l_n}$ будет один из членов многочлена $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$. Назовем *весом* этого члена число

$$l_1+2l_2+\ldots+nl_n,$$

т. е. сумму показателей, умноженных на индексы соответствующих σ_i. Это будет, иными словами, степень взятого нами члена по совокупности неизвестных $x_1, x_2, \ldots, x_n$, как вытекает из доказанной в § 51 теоремы о степени произведения многочленов. Тогда справедливо следующее утверждение:

Если однородный симметрический многочлен $f(x_1, x_2, \ldots, x_n)$ имеет по совокупности неизвестных степень s, то все члены его выражения $\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n)$ через σ будут одного и того же веса, равного s.

Действительно, если (2) из предшествующего параграфа есть высший член однородного многочлена f, то

$$s=k_1+k_2+\ldots+k_n.$$

Однако вес члена φ_1 равен, по (5) из предшествующего параграфа,

$$(k_1-k_2)+2(k_2-k_3)+\ldots+(n-1)(k_{n-1}-k_n)+nk_n=$$
$$=k_1+k_2+k_3+\ldots+k_n,$$

т. е. также равен s. Далее, многочлен $f_1=f-\varphi_1$ как разность двух однородных многочленов степени s сам будет однородным степени s, а поэтому и член φ_2 многочлена φ будет веса s и т. д.

Симметрические рациональные дроби. Основная теорема о симметрических многочленах может быть распространена на случай рациональных дробей. Назовем рациональную дробь $\frac{f}{g}$ от n неизвестных $x_1, x_2, \ldots, x_n$ *симметрической*, если она остается равной самой себе при любой перестановке неизвестных. Легко показать, что *это определение не зависит от того, берем ли мы дробь $\frac{f}{g}$ или равную ей дробь $\frac{f_0}{g_0}$*. Действительно, если ω есть некоторая перестановка наших неизвестных, а φ — произвольный многочлен от этих неизвестных, то условимся через φ^{ω} обозначать тот многочлен, в который переводится φ перестановкой ω. По предположению, при любом ω

$$\frac{f}{g}=\frac{f^{\omega}}{g^{\omega}},$$

т. е. $fg^{\omega}=gf^{\omega}$. С другой стороны, из

$$\frac{f}{g}=\frac{f_0}{g_0}$$

следует $fg_0=gf_0$, откуда $f^{\omega}g_0^{\omega}=g^{\omega}f_0^{\omega}$. Умножая обе части последнего равенства на f, мы получаем:

$$ff^{\omega}g_0^{\omega}=fg^{\omega}f_0^{\omega}=gf^{\omega}f_0^{\omega},$$

откуда после сокращения на f^{ω} следует: $fg_0^{\omega}=gf_0^{\omega}$, т. е.

$$\frac{f_0^{\omega}}{g_0^{\omega}}=\frac{f}{g}=\frac{f_0}{g_0}.$$

Справедлива следующая теорема:

Всякая симметрическая рациональная дробь от неизвестных $x_1, x_2, \ldots, x_n$ с коэффициентами из поля P представима в виде рациональной дроби от элементарных симметрических многочленов $\sigma_1, \sigma_2, \ldots, \sigma_n$ с коэффициентами, снова принадлежащими к P.

Действительно, пусть дана симметрическая рациональная дробь

$$\frac{f(x_1, x_2, \ldots, x_n)}{g(x_1, x_2, \ldots, x_n)}.$$

Предполагая ее несократимой, можно было бы доказать, что и f и g будут симметрическими многочленами. Следующий путь будет, однако, более простым. Если многочлен g не является симметрическим, то умножаем числитель и знаменатель на произведение всех $n!-1$ многочленов, получающихся из g при всевозможных нетождественных подстановках неизвестных. Легко проверить, что знаменатель будет теперь симметрическим многочленом. Ввиду симметричности всей дроби отсюда следует, что числитель теперь также будет симметрическим, а поэтому для доказательства теоремы остается выразить числитель и знаменатель через элементарные симметрические многочлены.

Степенные суммы. В приложениях часто встречаются симметрические многочлены

$$s_k=x_1^k+x_2^k+\ldots+x_n^k, \quad k=1, 2, \ldots,$$

т. е. суммы k-х степеней неизвестных $x_1, x_2, \ldots, x_n$. Эти многочлены, называемые *степенными суммами*, должны выражаться, по основной теореме, через элементарные симметрические многочлены. Разыскание этих выражений является, однако, при больших k весьма затруднительным, и поэтому представляет интерес та связь между многочленами $s_1, s_2, \ldots$ и $\sigma_1, \sigma_2, \ldots, \sigma_n$, которая будет сейчас установлена.

Прежде всего $s_1=\sigma_1$. Далее, если $k\leqslant n$, то легко проверить справедливость равенств

$$\left.\begin{aligned} &s_{k-1}\sigma_1=s_k+S\left(x_1^{k-1}x_2\right)^1), \\ &s_{k-2}\sigma_2=S\left(x_1^{k-1}x_2\right)+S\left(x_1^{k-2}x_2x_3\right), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &s_{k-i}\sigma_i=S\left(x_1^{k-i+1}x_2\dots x_i\right)+S\left(x_1^{k-i}x_2\dots x_ix_{i+1}\right), \\ &\qquad\qquad\qquad\qquad 2\leqslant i\leqslant k-2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &s_1\sigma_{k-1}=S\left(x_1^2x_2\dots x_{k-1}\right)+k\sigma_k. \end{aligned}\right\} \qquad (1)$$

Беря альтернированную сумму этих равенств (т. е. сумму с чередующимися знаками), а затем перенося все члены в одну часть равенства, мы получим следующую формулу:

$$s_k-s_{k-1}\sigma_1+s_{k-2}\sigma_2-\dots+(-1)^{k-1}s_1\sigma_{k-1}+(-1)^k k\sigma_k=0 \quad (2)$$
$$(k\leqslant n).$$

Если же $k>n$, то система равенств (1) примет вид

$$\begin{aligned} &s_{k-1}\sigma_1=s_k+S\left(x_1^{k-1}x_2\right), \\ &s_{k-2}\sigma_2=S\left(x_1^{k-1}x_2\right)+S\left(x_1^{k-2}x_2x_3\right), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &s_{k-i}\sigma_i=S\left(x_1^{k-i+1}x_2\dots x_i\right)+S\left(x_1^{k-i}x_2\dots x_ix_{i+1}\right), \qquad 2\leqslant i\leqslant n-1, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &s_{k-n}\sigma_n=S\left(x_1^{k-n+1}x_2\dots x_n\right), \end{aligned}$$

откуда вытекает формула

$$s_k-s_{k-1}\sigma_1+s_{k-2}\sigma_2-\dots+(-1)^n s_{k-n}\sigma_n=0 \qquad (k>n). \qquad (3)$$

Формулы (2) и (3) называются *формулами Ньютона*. Они связывают степенные суммы с элементарными симметрическими многочленами и позволяют последовательно находить выражения для s_1, s_2, s_3, ... через $\sigma_1, \sigma_2, \dots, \sigma_n$. Так, мы знаем, что $s_1=\sigma_1$, что вытекает и из формулы (2). Если, далее, $k=2\leqslant n$, то, по (2), $s_2-s_1\sigma_1+2\sigma_2=0$, откуда

$$s_2=\sigma_1^2-2\sigma_2.$$

Далее, при $k=3\leqslant n$ будет $s_3-s_2\sigma_1+s_1\sigma_2-3\sigma_3=0$, откуда, используя найденные уже выражения для s_1 и s_2, получаем:

$$s_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3,$$

---

[1]) См. (11) предшествующего параграфа.

что нам уже известно (см. (12) из предшествующего параграфа). Если же $k=3$, но $n=2$, то, по (3), $s_3-s_2\sigma_1+s_1\sigma_2=0$, откуда $s_3=\sigma_1^3-3\sigma_1\sigma_2$. Пользуясь формулами Ньютона, можно получить общую формулу, выражающую s_k через $\sigma_1, \sigma_2, \ldots, \sigma_n$. Эта формула, впрочем, весьма громоздка и мы не будем ее приводить.

Если основное поле P имеет характеристику 0 и поэтому деление на любое натуральное число n имеет смысл[1], то формула (2) дает возможность последовательно выразить элементарные симметрические многочлены $\sigma_1, \sigma_2, \ldots, \sigma_n$ через первые n степенных сумм $s_1, s_2, \ldots, s_n$. Так, $\sigma_1=s_1$, а поэтому

$$\sigma_2=\frac{1}{2}(s_1\sigma_1-s_2)=\frac{1}{2}(s_1^2-s_2),$$
$$\sigma_3=\frac{1}{3}(s_3-s_2\sigma_1+s_1\sigma_2)=\frac{1}{6}(s_1^3-3s_1s_2+2s_3)$$

и т. д. Отсюда и из основной теоремы вытекает следующий результат:

Всякий симметрический многочлен от n неизвестных $x_1, x_2, \ldots$ $\ldots, x_n$ над полем P характеристики нуль представим в виде многочлена от степенных сумм $s_1, s_2, \ldots, s_n$ с коэффициентами, принадлежащими к полю P.

Многочлены, симметрические по двум системам неизвестных. В следующем параграфе, а также в § 58 будет использовано одно обобщение понятия симметрического многочлена. Пусть даны две системы неизвестных $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_r$, причем их объединение

$$x_1, x_2, \ldots, x_n, \quad y_1, y_2, \ldots, y_r \tag{4}$$

алгебраически независимо над полем P. Многочлен над полем P $f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r)$ называется *симметрическим по двум системам неизвестных*, если он не меняется при любых перестановках неизвестных $x_1, x_2, \ldots, x_n$ между собой и неизвестных $y_1, y_2, \ldots, y_r$ между собой. Если для элементарных симметрических многочленов от $x_1, x_2, \ldots, x_n$ мы сохраним обозначения $\sigma_1, \sigma_2, \ldots, \sigma_n$, а элементарные симметрические многочлены от $y_1, y_2, \ldots, y_r$ обозначим через $\tau_1, \tau_2, \ldots, \tau_r$, то основная теорема обобщается следующим образом.

Всякий многочлен $f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r)$ над полем P, симметрический по системам неизвестных $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_r$, представим в виде многочлена (с коэффициентами из P) от элементарных симметрических многочленов по этим двум системам неизвестных:

$$f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_r)=\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r).$$

---

[1] В поле характеристики p выражение $\frac{a}{p}$ не имеет смысла при $a\neq 0$, так как в этом поле при любом x будет $px=0$.

В самом деле, многочлен f можно рассматривать как многочлен $\bar{f}(y_1, y_2, \ldots, y_r)$ с коэффициентами, являющимися многочленами от $x_1, x_2, \ldots, x_n$. Так как f не меняется при перестановках неизвестных $x_1, x_2, \ldots, x_n$, то коэффициенты многочлена $\bar{f}$ будут симметрическими многочленами от $x_1, x_2, \ldots, x_n$ и поэтому, по основной теореме, представимы в виде многочленов (с коэффициентами из P) от $\sigma_1, \sigma_2, \ldots, \sigma_n$. С другой стороны, многочлен $\bar{f}(y_1, y_2, \ldots, y_r)$, рассматриваемый над полем $P(x_1, x_2, \ldots, x_n)$, будет симметрическим относительно $y_1, y_2, \ldots, y_r$ и поэтому представим в виде многочлена $\bar{\varphi}(\tau_1, \tau_2, \ldots, \tau_r)$. Коэффициенты многочлена $\bar{\varphi}$ будут, как показано в начале настоящего параграфа, выражаться через коэффициенты многочлена $\bar{f}$ при помощи сложения и вычитания, а поэтому они также будут многочленами от $\sigma_1, \sigma_2, \ldots, \sigma_n$. Это приводит, очевидно, к искомому выражению для f через $\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r$.

Пример. Многочлен

$$f(x_1, x_2, x_3, y_1, y_2) = x_1x_2x_3 - x_1x_2y_1 - x_1x_2y_2 - x_1x_3y_1 - \\ - x_1x_3y_2 - x_2x_3y_1 - x_2x_3y_2 + x_1y_1y_2 + x_2y_1y_2 + x_3y_1y_2$$

симметричен как по неизвестным x_1, x_2, x_3, так и по неизвестным y_1, y_2, но не будет симметрическим по всей совокупности пяти неизвестных, как обнаруживается хотя бы при транспозиции неизвестных x_1 и y_1. Найдем выражение для f через $\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2$:

$$f = x_1x_2x_3 - (x_1x_2 + x_1x_3 + x_2x_3)\, y_1 - (x_1x_2 + x_1x_3 + x_2x_3)\, y_2 + \\ + (x_1 + x_2 + x_3)\, y_1y_2 = \sigma_3 - \sigma_2y_1 - \sigma_2y_2 + \sigma_1y_1y_2 = \sigma_3 - \sigma_2\tau_1 + \sigma_1\tau_2.$$

Доказанная сейчас теорема распространяется, понятно, также на случай трех и большего числа систем неизвестных.

Для многочленов, симметрических по двум системам неизвестных, справедлива также теорема единственности представления через элементарные симметрические многочлены. Иными словами, справедлива следующая теорема:

Объединенная система

$$\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r$$

элементарных симметрических многочленов от заданных систем неизвестных $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_r$ алгебраически независима над полем P.

Пусть, в самом деле, существует многочлен

$$\varphi(\sigma_1, \sigma_2, \ldots, \sigma_n, \tau_1, \tau_2, \ldots, \tau_r)$$

над полем P, равный нулю, хотя не все его коэффициенты нули. Этот многочлен можно рассматривать как многочлен $\psi(\tau_1\ \tau_2, \ldots, \tau_r)$ с коэффициентами, являющимися многочленами от $\sigma_1, \sigma_2, \ldots, \sigma_n$.

Можно считать, следовательно, что ψ — многочлен от $\tau_1, \tau_2, \ldots, \tau_r$ над полем рациональных дробей

$$Q = P(x_1, x_2, \ldots, x_n).$$

Система $y_1, y_2, \ldots, y_r$ остается алгебраически независимой над полем Q: если бы для этой системы существовала алгебраическая зависимость с коэффициентами из Q, то, освобождаясь от знаменателей, мы получили бы алгебраическую зависимость в системе (4) против предположения. Опираясь на теорему единственности из предыдущего параграфа, мы получаем теперь, что система $\tau_1, \tau_2, \ldots$ $\ldots, \tau_r$ также должна быть алгебраически независимой над полем Q, а поэтому все коэффициенты многочлена ψ равны нулю. Эти коэффициенты являются, однако, многочленами от $\sigma_1, \sigma_2, \ldots, \sigma_n$, а поэтому, снова на основании теоремы единственности для случая одной системы неизвестных (на этот раз системы $x_1, x_2, \ldots, x_n$), все коэффициенты этих последних многочленов сами равны нулю. Этим доказано, что в противоречие с предположением все коэффициенты многочлена φ должны быть равными нулю.

§ 54*. Результант. Исключение неизвестного. Дискриминант

Если дан многочлен $f(x_1, x_2, \ldots, x_n)$ из кольца $P[x_1, x_2, \ldots, x_n]$, то его *решением* называется такая система значений для неизвестных

$$x_1 = \alpha_1, \quad x_2 = \alpha_2, \ldots, \quad x_n = \alpha_n,$$

взятых в поле P или в некотором расширении $\bar{P}$ этого поля, которая обращает многочлен f в нуль:

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0.$$

Всякий многочлен f, степень которого больше нуля, обладает решениями: если неизвестное x_1 входит в запись этого многочлена, то в качестве $\alpha_2, \ldots, \alpha_n$ можно взять по существу произвольные элементы из поля P, лишь бы степень многочлена $f(x_1, \alpha_2, \ldots, \alpha_n)$ оставалась строго положительной, а затем, используя теорему о существовании корня (§ 49), взять такое расширение $\bar{P}$ поля P, в котором многочлен $f(x_1, \alpha_2, \ldots, \alpha_n)$ от одного неизвестного x_1 обладает корнем α_1. Мы видим вместе с тем, что свойство многочлена степени n от одного неизвестного обладать во всяком поле не более чем n корнями для многочленов от нескольких неизвестных перестает быть справедливым.

Если дано несколько многочленов от n неизвестных, то можно поставить вопрос о разыскании решений, общих для всех этих многочленов, т. е. решений той системы уравнений, которая получается в результате приравнивания заданных многочленов нулю. Частный случай этой задачи, а именно случай систем линейных уравнений,

уже был подвергнут во второй главе детальному рассмотрению. Однако для противоположного частного случая одного уравнения от одного неизвестного, но имеющего произвольную степень, мы не знаем о корнях ничего, кроме того, что они существуют в некотором расширении основного поля. Разыскание и изучение решений произвольной нелинейной системы уравнений от нескольких неизвестных является, понятно, еще более сложной задачей, выходящей, впрочем, за рамки нашего курса и составляющей предмет особой математической науки — алгебраической геометрии. Мы же здесь ограничимся лишь случаем системы двух уравнений произвольной степени от двух неизвестных и покажем, что этот случай может быть сведен к случаю одного уравнения от одного неизвестного.

Займемся сперва вопросом о существовании общих корней у двух многочленов от одного неизвестного. Пусть даны многочлены

$$\left.\begin{aligned} f(x) &= a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \\ g(x) &= b_0x^s + b_1x^{s-1} + \ldots + b_{s-1}x + b_s \end{aligned}\right\} \qquad (1)$$

над полем P, причем $a_0 \neq 0$, $b_0 \neq 0$.

Из результатов предшествующей главы без труда вытекает, что *многочлены $f(x)$ и $g(x)$ тогда и только тогда обладают общим корнем в некотором расширении поля P, если они не являются взаимно простыми.* Таким образом, вопрос о существовании общих корней у данных многочленов может быть решен применением к ним алгоритма Евклида.

Сейчас мы укажем другой метод для получения ответа на этот вопрос. Пусть $\overline{P}$ будет некоторое такое расширение поля P, в котором $f(x)$ имеет n корней $\alpha_1, \alpha_2, \ldots, \alpha_n$, а $g(x)$ имеет s корней $\beta_1, \beta_2, \ldots, \beta_s$; в качестве $\overline{P}$ можно взять поле разложения для произведения $f(x)g(x)$. Элемент

$$R(f, g) = a_0^s b_0^n \prod_{i=1}^{n} \prod_{j=1}^{s} (\alpha_i - \beta_j) \qquad (2)$$

поля $\overline{P}$ называется *результантом* многочленов $f(x)$ и $g(x)$. Очевидно, что *$f(x)$ и $g(x)$ тогда и только тогда обладают в $\overline{P}$ общим корнем, если $R(f, g) = 0$.* Так как

$$g(x) = b_0 \prod_{j=1}^{s} (x - \beta_j)$$

и поэтому

$$g(\alpha_i) = b_0 \prod_{j=1}^{s} (\alpha_i - \beta_j),$$

то результант $R(f, g)$ может быть записан также в виде

$$R(f, g) = a_0^s \prod_{i=1}^{n} g(\alpha_i). \qquad (3)$$

Многочлены $f(x)$ и $g(x)$ используются в определении результанта не симметричным образом. Действительно,

$$R(g, f)=b_0^n a_0^s \prod_{j=1}^{s} \prod_{i=1}^{n} (\beta_j-\alpha_i)=(-1)^{ns} R(f; g). \tag{4}$$

В соответствии с (3) $R(g, f)$ можно записать в виде

$$R(g, f)=b_0^n \prod_{j=1}^{s} f(\beta_j). \tag{5}$$

Выражение (2) для результанта требует знания корней многочленов $f(x)$ и $g(x)$ и поэтому практически бесполезно для решения вопроса о существовании у этих двух многочленов общего корня. Оказывается, однако, что *результант $R(f, g)$ может быть представлен в виде многочлена от коэффициентов $a_0, a_1, \ldots, a_n$, $b_0, b_1, \ldots, b_s$ многочленов $f(x)$ и $g(x)$.*

Возможность такого представления легко вытекает из результатов предшествующего параграфа. В самом деле, формула (2) показывает, что результант $R(f, g)$ является симметрическим многочленом от двух систем неизвестных: системы $\alpha_1, \alpha_2, \ldots, \alpha_n$ и системы $\beta_1, \beta_2, \ldots, \beta_s$. Он представим поэтому, как доказано в конце предшествующего параграфа, в виде многочлена от элементарных симметрических многочленов по этим двум системам неизвестных, т. е., ввиду формул Вьета, в виде многочлена от частных $\frac{a_i}{a_0}$, $i=1, 2, \ldots, n$, и $\frac{b_j}{b_0}$, $j=1, 2, \ldots, s$; множитель $a_0^s b_0^n$, включенный в (2), освобождает полученное выражение от a_0 и b_0 в знаменателях. Впрочем, было бы затруднительным разыскивать выражение результанта через коэффициенты при помощи методов, изложенных в предшествующих параграфах, и мы воспользуемся иным приемом.

Выражение для результанта многочленов (1), которое мы найдем, будет годно для любой пары таких многочленов. Мы будем считать, точнее говоря, что система корней

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_s \tag{6}$$

многочленов (1) является системой $n+s$ независимых неизвестных, т. е. системой $n+s$ элементов, алгебраически независимых над полем P в смысле § 51.

Мы получим выражение для результанта, которое, рассматриваемое как многочлен от неизвестных (6) (после замены по формулам Вьета коэффициентов через корни), будет равно правой части равенства (2), также рассматриваемой как многочлен от неизвестных (6).

Понимая равенство именно в смысле такого тождественного равенства относительно системы неизвестных (6), мы докажем, что

результант $R(f, g)$ *многочленов* (1) *равен следующему определителю порядка* $n+s$:

$$D=\left|\begin{array}{ccccccc} a_0 & a_1 & \dots & a_n & & & \\ & a_0 & a_1 & \dots & a_n & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & a_0 & a_1 & \dots & a_n \\ b_0 & b_1 & \dots & b_s & & & \\ & b_0 & b_1 & \dots & b_s & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b_0 & b_1 & \dots & b_s \end{array}\right| \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} s \text{ строк} \\ \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} n \text{ строк} \end{array} \qquad (7)$$

(на свободных местах стоят нули). Строение этого определителя достаточно ясно; мы отметим лишь, что на его главной диагонали стоит s раз коэффициент a_0 и затем n раз коэффициент b_s.

Для доказательства нашего утверждения мы двумя способами вычислим произведение $a_0^s b_0^n DM$, где M есть следующий вспомогательный определитель порядка $n+s$:

$$M=\left|\begin{array}{cccccccc} \beta_1^{n+s-1} & \beta_2^{n+s-1} & \dots & \beta_s^{n+s-1} & \alpha_1^{n+s-1} & \alpha_2^{n+s-1} & \dots & \alpha_n^{n+s-1} \\ \beta_1^{n+s-2} & \beta_2^{n+s-2} & \dots & \beta_s^{n+s-2} & \alpha_1^{n+s-2} & \alpha_2^{n+s-2} & \dots & \alpha_n^{n+s-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \beta_1^2 & \beta_2^2 & \dots & \beta_s^2 & \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \beta_1 & \beta_2 & \dots & \beta_s & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \end{array}\right|.$$

M является определителем Вандермонда и поэтому равен, как указано в § 6, произведению разностей элементов его предпоследней строки, причем из всякого предшествующего элемента вычитается любой следующий элемент. Таким образом,

$$M=\prod_{1\leqslant i<j\leqslant s}(\beta_i-\beta_j)\cdot\prod_{j=1}^{s}\prod_{i=1}^{n}(\beta_j-\alpha_i)\cdot\prod_{1\leqslant i<j\leqslant n}(\alpha_i-\alpha_j)$$

и поэтому, ввиду (4),

$$a_0^s b_0^n DM=D\cdot \dot{R}(g, f)\cdot\prod_{1\leqslant i<j\leqslant s}(\beta_i-\beta_j)\cdot\prod_{1\leqslant i<j\leqslant n}(\alpha_i-\alpha_j). \qquad (8)$$

Вычислим, с другой стороны, произведение DM на основании теоремы об определителе произведения матриц. Перемножая соответствующие матрицы и учитывая, что все α являются корнями

для $f(x)$, а все β — корнями для $g(x)$, мы получим:

$$DM =$$

$$= \begin{vmatrix} \beta_1^{s-1} f(\beta_1) & \beta_2^{s-1} f(\beta_2) & \dots & \beta_s^{s-1} f(\beta_s) & 0 & 0 & \dots & 0 \\ \beta_1^{s-2} f(\beta_1) & \beta_2^{s-2} f(\beta_2) & \dots & \beta_s^{s-2} f(\beta_s) & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \beta_1 f(\beta_1) & \beta_2 f(\beta_2) & \dots & \beta_s f(\beta_s) & 0 & 0 & \dots & 0 \\ f(\beta_1) & f(\beta_2) & \dots & f(\beta_s) & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_1^{n-1} g(\alpha_1) & \alpha_2^{n-1} g(\alpha_2) & \dots & \alpha_n^{n-1} g(\alpha_n) \\ 0 & 0 & \dots & 0 & \alpha_1^{n-2} g(\alpha_1) & \alpha_2^{n-2} g(\alpha_2) & \dots & \alpha_n^{n-2} g(\alpha_n) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \alpha_1 g(\alpha_1) & \alpha_2 g(\alpha_2) & \dots & \alpha_n g(\alpha_n) \\ 0 & 0 & \dots & 0 & g(\alpha_1) & g(\alpha_2) & \dots & g(\alpha_n) \end{vmatrix}.$$

Применяя теорему Лапласа, вынося затем общие множители из столбцов определителей и вычисляя остающиеся определители как определители Вандермонда, мы получим:

$$a_0^s b_0^n DM = a_0^s b_0^n \prod_{j=1}^{s} f(\beta_j) \cdot \prod_{1 \leqslant i < j \leqslant s} (\beta_i - \beta_j) \cdot \prod_{i=1}^{n} g(\alpha_i) \cdot \prod_{1 \leqslant i < j \leqslant n} (\alpha_i - \alpha_j)$$

или, используя (3) и (5),

$$a_0^s b_0^n DM = R(f, g) R(g, f) \cdot \prod_{1 \leqslant i < j \leqslant s} (\beta_i - \beta_j) \cdot \prod_{1 \leqslant i < j \leqslant n} (\alpha_i - \alpha_j). \quad (9)$$

Мы получаем, что правые части равенств (8) и (9), рассматриваемые как многочлены от неизвестных (6), равны между собой. Обе части полученного равенства можно сократить на общие множители, не равные тождественно нулю. Общий множитель $R(g, f)$ не равен нулю: так как $a_0 \neq 0$ и $b_0 \neq 0$ по условию, то достаточно подобрать для неизвестных (6) не равные друг другу значения (в основном поле или в некотором его расширении), чтобы из (4) получить отличное от нуля значение для многочлена $R(g, f)$. Так же доказывается, что и другие два общих множителя отличны от нуля. Сокращая на все эти общие множители, мы приходим к равенству

$$R(f, g) = D, \quad (10)$$

которое и требовалось доказать.

Откажемся теперь от требования, чтобы старшие коэффициенты многочленов (1) были отличны от нуля[1]). Об истинных степенях этих многочленов, можно, следова-

[1]) Этот временный отказ от того условия о старшем коэффициенте многочлена, которому мы следовали до сих пор, обусловлен дальнейшими приложениями: мы хотим рассматривать системы многочленов от двух неизвестных и будем одно из этих неизвестных относить в коэффициенты. Старший коэффициент может, следовательно, обратиться в нуль при частных значениях этого неизвестного.

тельно, лишь утверждать, что они не больше их «формальных» степеней n и, соответственно, s. Выражение (2) для результанта не имеет теперь смысла, так как рассматриваемые многочлены имеют, возможно, меньше корней, чем n или s. С другой стороны, определитель (7) и теперь может быть написан, и так как уже доказано, что при $a_0 \neq 0$, $b_0 \neq 0$ этот определитель равен результанту, то и в нашем общем случае назовем его *результантом* многочленов $f(x)$ и $g(x)$ и обозначим через $R(f, g)$.

Теперь уже нельзя, однако, рассчитывать на то, что равенство результанта нулю равносильно существованию у наших многочленов общего корня. Действительно, если $a_0 = 0$ и $b_0 = 0$, то $R(f, g) = 0$ независимо от того, обладают ли многочлены f и g общими корнями или нет. Оказывается, однако, что этот случай будет единственным, когда из равенства результанта нулю нельзя вывести заключение о существовании у данных многочленов общих корней[1]). Именно, справедлива следующая теорема:

Если даны многочлены (1) *с произвольными старшими коэффициентами, то результант* (7) *этих многочленов тогда и только тогда равен нулю, если эти многочлены обладают общим корнем или же если их старшие коэффициенты оба равны нулю.*

Доказательство. Случай $a_0 \neq 0$, $b_0 \neq 0$ уже рассматривался выше, а случай $a_0 = b_0 = 0$ предусмотрен в формулировке теоремы. Нам остается рассмотреть случай, когда один из старших коэффициентов многочленов (1), например a_0, отличен от нуля, а b_0 равно нулю.

Если $b_i = 0$ для всех i, $i = 0, 1, \ldots, s$, то $R(f, g) = 0$, так как определитель (7) содержит нулевые строки. В этом случае, однако, многочлен $g(x)$ равен нулю тождественно и поэтому имеет общие корни с $f(x)$. Если же

$$b_0 = b_1 = \ldots = b_{k-1} = 0, \quad \text{но} \quad b_k \neq 0, \quad k \leqslant s,$$

и если

$$\bar{g}(x) = b_k x^{s-k} + b_{k+1} x^{s-k-1} + \ldots + b_{s-1} x + b_s,$$

то, заменяя в определителе (7) элементы $b_0, b_1, \ldots, b_{k-1}$ нулями и применяя теорему Лапласа, мы придем, очевидно, к равенству

$$R(f, g) = a_0^k R(f, \bar{g}). \tag{11}$$

Так как, однако, старшие коэффициенты обоих многочленов f и $\bar{g}$ отличны от нуля, то, по доказанному выше, равенство $R(f, \bar{g}) = 0$ необходимо и достаточно для существования общего корня у многочленов f и $\bar{g}$. С другой стороны, по (11), равенства $R(f, g) = 0$ и $R(f, \bar{g}) = 0$ равносильны, а так как многочлены g и $\bar{g}$ имеют,

[1]) Определитель (7) равен нулю, конечно, и при $a_n = b_s = 0$. В этом случае, однако, многочлены (1) имеют общий корень 0.

понятно, одинаковые корни, то мы получаем, что и в рассматриваемом случае равенство нулю результанта $R(f, g)$ равносильно существованию общего корня у многочленов $f(x)$ и $g(x)$. Этим теорема доказана.

Найдем результант двух квадратных многочленов

$$f(x)=a_0x^2+a_1x+a_2, \quad g(x)=b_0x^2+b_1x+b_2.$$

По (7)

$$R(f, g)=\begin{vmatrix} a_0 & a_1 & a_2 & 0 \\ 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & 0 \\ 0 & b_0 & b_1 & b_2 \end{vmatrix},$$

или, вычисляя определитель разложением по первой и третьей строкам,

$$R(f, g)=(a_0b_2-a_2b_0)^2-(a_0b_1-a_1b_0)(a_1b_2-a_2b_1). \qquad (12)$$

Так, если даны многочлены

$$f(x)=x^2-6x+2, \quad g(x)=x^2+x+5,$$

то, по (12), $R(f, g)=233$, и потому эти многочлены не имеют общих корней. Если же даны многочлены

$$f(x)=x^2-4x-5, \quad g(x)=x^2-7x+10,$$

то $R(f, g)=0$, т. е. эти многочлены обладают общим корнем; этим корнем является число 5.

Исключение неизвестного из системы двух уравнений с двумя неизвестными. Пусть даны два многочлена f и g от двух неизвестных x и y с коэффициентами из некоторого поля P. Мы запишем эти многочлены по убывающим степеням неизвестного x:

$$\left.\begin{aligned} f(x, y)&=a_0(y)x^k+a_1(y)x^{k-1}+\ldots+a_{k-1}(y)x+a_k(y), \\ g(x, y)&=b_0(y)x^l+b_1(y)x^{l-1}+\ldots+b_{l-1}(y)x+b_l(y); \end{aligned}\right\} \qquad (13)$$

коэффициенты будут многочленами из кольца $P[y]$. Найдем результант многочленов f и g, рассматриваемых как многочлены от x, и обозначим его через $R_x(f, g)$; он будет, ввиду (7), многочленом от одного неизвестного y с коэффициентами из поля P:

$$R_x(f, g)=F(y). \qquad (14)$$

Пусть система многочленов (13) обладает в некотором расширении поля P общим решением $x=\alpha$, $y=\beta$. Подставляя в (13) вместо y значение β, мы получим два многочлена $f(x, \beta)$ и $g(x, \beta)$ от одного неизвестного x. Эти многочлены обладают общим корнем α, а поэтому их результант, равный, ввиду (14), $F(\beta)$, должен быть равным нулю, т. е. β *должно быть корнем результанта* $R_x(f, g)$. Обратно, если результант $R_x(f, g)$ многочленов (13) обладает корнем β, то результант многочленов $f(x, \beta)$ и $g(x, \beta)$ равен нулю, т. е. *либо*

эти многочлены обладают общим корнем, либо же оба их старших коэффициента равны нулю,

$$a_0(\beta)=b_0(\beta)=0.$$

Этим путем разыскание общих решений системы многочленов (13) сведено к разысканию корней одного многочлена (14) от одного неизвестного y, т. е., как принято говорить, ***неизвестное x исключено из системы многочленов*** (13).

Следующая теорема отвечает на вопрос о степени того многочлена, который мы получаем после исключения одного неизвестного из системы двух многочленов с двумя неизвестными:

Если многочлены $f(x, y)$ и $g(x, y)$ имеют по совокупности неизвестных соответственно степени n и s, то степень многочлена $R_x(f, g)$ по неизвестному y не больше произведения ns, если, конечно, этот многочлен не равен нулю тождественно.

Прежде всего, если мы рассматриваем два многочлена от одного неизвестного со старшими коэффициентами, равными единице, то, по (2), их результант $R(f, g)$ является однородным многочленом от $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_s$ степени ns. Отсюда следует, что если в выражение результанта через коэффициенты $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_s$ входит член

$$a_1^{k_1}a_2^{k_2}\ldots a_n^{k_n}b_1^{l_1}b_2^{l_2}\ldots b_s^{l_s}$$

и если *весом* этого члена будет названо число

$$k_1+2k_2+\ldots+nk_n+l_1+2l_2+\ldots+sl_s,$$

то *все члены выражения $R(f, g)$ через коэффициенты имеют один и тот же вес, равный ns.* Это утверждение справедливо и в общем случае, для членов результанта (7), если *весом* члена $a_0^{k_0}a_1^{k_1}\ldots a_n^{k_n}b_0^{l_0}b_1^{l_1}\ldots b_s^{l_s}$ будет названо число

$$0\cdot k_0+1\cdot k_1+\ldots+nk_n+0\cdot l_0+1\cdot l_1+\ldots+sl_s. \tag{15}$$

Действительно, заменяя в членах определителя (7) множители a_0 и b_0 единицей, мы приходим к уже рассмотренному случаю, однако показатели при этих множителях входят в (15) с коэффициентами 0.

Запишем теперь многочлены f и g в следующем виде:

$$f(x, y)=a_0(y)x^n+a_1(y)x^{n-1}+\ldots+a_n(y),$$
$$g(x, y)=b_0(y)x^s+b_1(y)x^{s-1}+\ldots+b_s(y).$$

Так как n есть степень $f(x, y)$ по совокупности неизвестных, то степень коэффициента $a_r(y)$, $r=0, 1, 2, \ldots, n$, не может превосходить его индекс r; это же верно и для $b_r(y)$. Отсюда следует, что степень каждого члена результанта $R_x(f, g)$ не больше веса этого члена, т. е. она не больше числа ns, что и требовалось доказать.

Примеры.

1. Найти общие решения системы многочленов

$$f(x, y)=x^2y+3xy+2y+3,$$
$$g(x, y)=2xy-2x+2y+3.$$

Исключим из этой системы неизвестное x, для чего перепишем ее в виде

$$\left.\begin{aligned} f(x, y) &= y\cdot x^2 + (3y)\cdot x + (2y+3), \\ g(x, y) &= (2y-2)x + (2y+3); \end{aligned}\right\} \tag{16}$$

тогда

$$R_x(f, g) = \begin{vmatrix} y & 3y & 2y+3 \\ 2y-2 & 2y+3 & 0 \\ 0 & 2y-2 & 2y+3 \end{vmatrix} = 2y^2 + 11y + 12.$$

Корнями результанта будут числа $\beta_1 = -4$, $\beta_2 = -\frac{3}{2}$. При этих значениях неизвестного y старшие коэффициенты многочленов (16) не обращаются в нуль, поэтому каждое из них вместе с некоторым значением для x составляет решение заданной системы многочленов. Многочлены

$$f(x, -4) = -4x^2 - 12x - 5,$$
$$g(x, -4) = -10x - 5$$

обладают общим корнем $\alpha_1 = -\frac{1}{2}$. Многочлены

$$f\left(x, -\frac{3}{2}\right) = -\frac{3}{2}x^2 - \frac{9}{2}x,$$
$$g\left(x, -\frac{3}{2}\right) = -5x$$

имеют общий корень $\alpha_2 = 0$. Таким образом, заданная система многочленов имеет два решения:

$$\alpha_1 = -\frac{1}{2}, \quad \beta_1 = -4 \quad \text{и} \quad \alpha_2 = 0, \quad \beta_2 = -\frac{3}{2}.$$

2. Исключить одно неизвестное из системы многочленов

$$f(x, y) = 2x^3y - xy^2 + x + 5,$$
$$g(x, y) = x^2y^2 + 2xy^2 - 5y + 1.$$

Так как оба многочлена имеют по неизвестному y степень 2, тогда как у одного из них по неизвестному x степень 3, то целесообразно исключить y. Перепишем систему в виде

$$\left.\begin{aligned} f(x, y) &= (-x)\cdot y^2 + (2x^3)\cdot y + (x+5), \\ g(x, y) &= (x^2+2x)y^2 - 5y + 1 \end{aligned}\right\} \tag{17}$$

и найдем ее результант, применяя формулу (12):

$$\begin{aligned} R_y(f, g) = {} & [(-x)\cdot 1 - (x+5)(x^2+2x)]^2 - \\ & - [(-x)(-5) - 2x^3(x^2+2x)][2x^3\cdot 1 - (x+5)(-5)] = \\ & = 4x^8 + 8x^7 + 11x^6 + 84x^5 + 161x^4 + 154x^3 + 96x^2 - 125x. \end{aligned}$$

Одним из корней результанта является 0. Однако при этом значении неизвестного x оба старших коэффициента многочленов (17) обращаются в нуль, причем, как легко видеть, многочлены $f(0, y)$ и $g(0, y)$ не имеют общих корней. У нас нет способа найти другие корни результанта. Можно утверждать лишь, что если бы мы их нашли (например, в поле разложения для $R_y(f, g)$), то ни один из них не обращал бы в нуль оба старших коэффициента многочленов (17) и поэтому каждый из этих корней вместе с некоторым значением для y (одним или даже несколькими) составлял бы решение заданной системы многочленов.

Существуют методы, позволяющие последовательно исключать неизвестные и из систем с произвольным числом многочленов и неизвестных. Эти методы, однако, слишком громоздки и поэтому не могут быть включены в наш курс.

Дискриминант. По аналогии с вопросом, который привел нас к понятию результанта, можно поставить вопрос об условиях, при которых многочлен $f(x)$ степени n из кольца $P[x]$ обладает кратными корнями. Пусть

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,\qquad a_0\neq 0,$$

и пусть в некотором расширении поля P этот многочлен имеет корни $\alpha_1, \alpha_2, \ldots, \alpha_n$. Очевидно, что *среди этих корней тогда и только тогда будут равные, если равно нулю произведение*

$$\begin{aligned}\Delta=(\alpha_2-\alpha_1)(\alpha_3-\alpha_1)\ \ldots\ (\alpha_n-\alpha_1)\times\\ \times(\alpha_3-\alpha_2)(\alpha_4-\alpha_2)\ \ldots\ (\alpha_n-\alpha_2)\times\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ \times(\alpha_n-\alpha_{n-1})=\prod_{n\geqslant i>j\geqslant 1}(\alpha_i-\alpha_j)\end{aligned}$$

или, что то же, если равно нулю произведение

$$D=a_0^{2n-2}\prod_{n\geqslant i>j\geqslant 1}(\alpha_i-\alpha_j)^2,$$

называемое *дискриминантом* многочлена $f(x)$.

В отличие от произведения Δ, могущего менять знак при перестановке корней, дискриминант D симметричен относительно $\alpha_1, \alpha_2, \ldots, \alpha_n$ и поэтому может быть выражен через коэффициенты многочлена $f(x)$. Для разыскания этого выражения в предположении, что поле P имеет характеристику нуль, можно воспользоваться связью, существующей между дискриминантом многочлена $f(x)$ и результантом этого многочлена и его производной. Наличие такой связи естественно ожидать: мы знаем из § 49, что многочлен тогда и только тогда обладает кратными корнями, если у него есть общие корни с производной $f'(x)$, а поэтому тогда и только тогда $D=0$, если $R(f, f')=0$.

По формуле (3) настоящего параграфа

$$R(f, f')=a_0^{n-1}\prod_{i=1}^{n}f'(\alpha_i).$$

Дифференцируя равенство

$$f(x)=a_0\prod_{k=1}^{n}(x-\alpha_k),$$

мы получаем:

$$f'(x)=a_0\sum_{k=1}^{n}\prod_{j\neq k}(x-\alpha_j).$$

После подстановки сюда α_i вместо x все слагаемые, кроме i-го, обращаются в нуль и поэтому

$$f'(\alpha_i) = a_0 \prod_{j \neq i} (\alpha_i - \alpha_j),$$

откуда

$$R(f, f') = a_0^{n-1} \cdot a_0^n \prod_{i=1}^{n} \prod_{j \neq i} (\alpha_i - \alpha_j).$$

В это произведение для любых i и j, $i > j$, входят два множителя: $\alpha_i - \alpha_j$ и $\alpha_j - \alpha_i$. Их произведение равно $(-1) \cdot (\alpha_i - \alpha_j)^2$, а так как существует $\frac{n(n-1)}{2}$ пар индексов i, j, удовлетворяющих неравенствам $n \geqslant i > j \geqslant 1$, то

$$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_0^{2n-1} \prod_{n \geqslant i > j \geqslant 1} (\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} a_0 D.$$

Пример. Найдем дискриминант квадратного трехчлена

$$f(x) = ax^2 + bx + c.$$

Так как $f'(x) = 2ax + b$, то

$$R(f, f') = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = a(-b^2 + 4ac).$$

В нашем случае $\frac{n(n-1)}{2} = 1$ и поэтому

$$D = -a^{-1} R(f, f') = b^2 - 4ac.$$

Это совпадает с тем, что в школьной алгебре называют обычно дискриминантом квадратного уравнения.

Другой способ разыскания дискриминанта состоит в следующем. Составим определитель Вандермонда из степеней корней $\alpha_1, \alpha_2, \ldots, \alpha_n$. Как доказано в § 6,

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \ldots & \alpha_n^{n-1} \end{vmatrix} = \prod_{n \geqslant i > j \geqslant 1} (\alpha_i - \alpha_j) = \Delta,$$

а поэтому дискриминант равен квадрату этого определителя, умножённому на a_0^{2n-2}. Умножая этот определитель на его транспониро-

ванный по правилу умножения матриц и вспоминая определенные в предыдущем параграфе степенные суммы, мы получим:

$$D=a_0^{2n-2}\begin{vmatrix} n & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{vmatrix}, \qquad (18)$$

где s_k есть сумма k-х степеней корней $\alpha_1, \alpha_2, \dots, \alpha_n$.

Пример. Найдем дискриминант кубичного многочлена $f(x)=x^3+ax^2+bx+c$. По (18)

$$D=\begin{vmatrix} 3 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}.$$

Как мы знаем из предыдущего параграфа,

$$s_1=\sigma_1=-a,$$
$$s_2=\sigma_1^2-2\sigma_2=a^2-2b,$$
$$s_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3=-a^3+3ab-3c.$$

Пользуясь формулой Ньютона, мы найдем также, ввиду $\sigma_4=0$, что

$$s_4=\sigma_1^4-4\sigma_1^2\sigma_2+4\sigma_1\sigma_3+2\sigma_2^2=a^4-4a^2b+4ac+2b^2.$$

Отсюда

$$D=3s_2s_4+2s_1s_2s_3-s_2^3-s_1^2s_4-3s_3^2=$$
$$=a^2b^2-4b^3-4a^3c+18abc-27c^2. \qquad (19)$$

В частности, при $a=0$, т. е. для неполного кубичного многочлена, мы получаем

$$D=-4b^3-27c^2$$

в полном соответствии с тем, что было сказано в § 38.

§ 55*. Второе доказательство основной теоремы алгебры комплексных чисел

Доказательство основной теоремы, приведенное в § 23, было совершенно неалгебраическим. Мы хотим изложить сейчас другое доказательство, использующее большой алгебраический аппарат — так, в нем существенно используется основная теорема о симметрических многочленах (§ 52), а также теорема о существовании поля разложения для всякого многочлена (§ 49),— в то время как неалгебраическая часть этого доказательства является минимальной и сведена к одному весьма простому утверждению.

Заметим сначала, что в § 23 доказана лемма о модуле старшего члена многочлена. Считая коэффициенты многочлена $f(x)$ действи-

тельными и полагая $k=1$, мы получаем из этой леммы такое следствие:

При действительных значениях x, достаточно больших по абсолютной величине, знак многочлена $f(x)$ с действительными коэффициентами совпадает со знаком его старшего члена.

Отсюда вытекает следующий результат:

Многочлен нечетной степени с действительными коэффициентами имеет хотя бы один действительный корень.

В самом деле, пусть

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n,$$

причем все коэффициенты действительны. Ввиду нечетности n старший член a_0x^n имеет при положительных и отрицательных значениях x разные знаки, а потому, как доказано выше, при положительных и отрицательных значениях x, достаточно больших по абсолютной величине, многочлен $f(x)$ также будет иметь разные знаки. Существуют, следовательно, такие действительные значения x, например a и b, что

$$f(a)<0, \quad f(b)>0.$$

Из курса анализа известно, однако, что многочлен (т. е. целая рациональная функция) $f(x)$ является функцией непрерывной, а поэтому, ввиду одного из основных свойств непрерывных функций, при некоторых действительных значениях x, заключенных между a и b, $f(x)$ принимает любое заданное значение, промежуточное между $f(a)$ и $f(b)$. Существует, в частности, такое α, лежащее между a и b, что $f(\alpha)=0$.

Опираясь на этот результат, мы докажем теперь следующее утверждение:

Всякий многочлен произвольной степени с действительными коэффициентами имеет хотя бы один комплексный корень.

Пусть, в самом деле, дан многочлен $f(x)$ с действительными коэффициентами, имеющий степень $n=2^kq$, где q — нечетное число. Так как случай $k=0$ уже рассмотрен выше, мы будем полагать $k>0$, т. е. считать n четным числом, и будем вести доказательство индукцией по k, предполагая, что наше утверждение уже доказано для всех многочленов с действительными коэффициентами, степень которых делится на 2^{k-1}, но не делится на 2^k[1]).

Пусть P будет полем разложения для многочлена $f(x)$ над полем комплексных чисел (см. § 49) и пусть $\alpha_1, \alpha_2, \ldots, \alpha_n$ будут корни $f(x)$, содержащиеся в поле P. Выберем произвольное действительное число c и возьмем элементы поля P, имеющие вид

$$\beta_{ij}=\alpha_i\alpha_j+c(\alpha_i+\alpha_j), \quad i<j. \qquad (1)$$

---

[1]) Эта степень может, следовательно, быть даже больше n.

Число элементов β_{ij} равно, очевидно,

$$\frac{n(n-1)}{2}=\frac{2^kq(2^kq-1)}{2}=2^{k-1}q(2^kq-1)=2^{k-1}q', \tag{2}$$

где q' есть нечетное число.

Построим теперь многочлен $g(x)$ из кольца $P[x]$, имеющий своими корнями все эти элементы β_{ij} и только их:

$$g(x)=\prod_{i,\ j,\ i<j}(x-\beta_{ij}).$$

Коэффициенты этого многочлена являются элементарными симметрическими многочленами от β_{ij}. Они будут, следовательно, ввиду (1), многочленами от $\alpha_1, \alpha_2, \ldots, \alpha_n$ с действительными коэффициентами (так как число c действительное), причем даже симметрическими многочленами. В самом деле, транспозиция любых двух α, например α_k и α_l, влечет за собой лишь перестановку в системе всех β_{ij}: всякое β_{kj}, где j отлично от k и от l, превращается в β_{lj} и обратно, в то время как β_{kl} и все β_{ij} при i и j, отличных от k и l, остаются на месте. Однако коэффициенты многочлена $g(x)$ не меняются при перестановке его корней.

Отсюда следует, ввиду основной теоремы о симметрических многочленах, что коэффициенты многочлена $g(x)$ будут многочленами (с действительными коэффициентами) от коэффициентов заданного многочлена $f(x)$ и поэтому сами будут действительными числами. Степень этого многочлена, равная числу корней β_{ij}, делится, по (2), на 2^{k-1}, но не делится на 2^k. Поэтому, по предположению индукции, хотя бы один из корней β_{ij} многочлена $g(x)$ должен быть комплексным числом.

Таким образом, при всяком выборе действительного числа c можно указать такую пару индексов i, j, где $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, что элемент $\alpha_i\alpha_j+c(\alpha_i+\alpha_j)$ является комплексным числом — напомним, что поле P содержит поле комплексных чисел в качестве подполя. Понятно, что при другом выборе числа c ему будет соответствовать в указанном смысле, вообще говоря, другая пара индексов. Однако существует бесконечно много различных действительных чисел c, в то время как в нашем распоряжении находится лишь конечное число различных пар i, j. Отсюда следует, что можно выбрать такие два *различных* действительных числа c_1 и c_2, $c_1 \neq c_2$, что им соответствует *одна и та же* пара индексов i, j, для которых

$$\left.\begin{aligned} \alpha_i\alpha_j+c_1(\alpha_i+\alpha_j)&=a, \\ \alpha_i\alpha_j+c_2(\alpha_i+\alpha_j)&=b \end{aligned}\right\} \tag{3}$$

являются комплексными числами.

Из равенств (3) вытекает:

$$(c_1 - c_2)(\alpha_i + \alpha_j) = a - b,$$

откуда следует:

$$\alpha_i + \alpha_j = \frac{a-b}{c_1 - c_2},$$

т. е. эта сумма оказывается комплексным числом. Отсюда и хотя бы из первого из равенств (3) следует, что произведение $\alpha_i\alpha_j$ также будет комплексным числом. Таким образом, элементы α_i и α_j оказываются корнями квадратного уравнения

$$x^2 - (\alpha_i + \alpha_j)x + \alpha_i\alpha_j = 0$$

с комплексными коэффициентами и поэтому, как вытекает из формулы для решения квадратного уравнения с комплексными коэффициентами, выведенной в § 38, они сами должны быть комплексными числами. Мы нашли, следовательно, среди корней многочлена $f(x)$ даже два комплексных корня и этим доказали наше утверждение.

Для полного доказательства основной теоремы остается рассмотреть случай многочлена с произвольными комплексными коэффициентами. Пусть

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

будет такой многочлен. Возьмем многочлен

$$\overline{f}(x) = \overline{a}_0x^n + \overline{a}_1x^{n-1} + \ldots + \overline{a}_n,$$

полученный из $f(x)$ заменой всех коэффициентов сопряженными комплексными числами, и рассмотрим произведение

$$F(x) = f(x)\overline{f}(x) = b_0x^{2n} + b_1x^{2n-1} + \ldots + b_kx^{2n-k} + \ldots + b_{2n},$$

где, очевидно,

$$b_k = \sum_{i+j=k} a_i\overline{a}_j, \qquad k = 0, 1, 2, \ldots, 2n.$$

Опираясь на известные нам из § 18 свойства сопряженных комплексных чисел, мы получаем, что

$$\overline{b}_k = \sum_{i+j=k} \overline{a}_i a_j = b_k,$$

т. е. все коэффициенты многочлена $F(x)$ оказываются действительными.

Отсюда, как доказано выше, следует, что многочлен $F(x)$ обладает хотя бы одним комплексным корнем β,

$$F(\beta) = f(\beta)\overline{f}(\beta) = 0,$$

т. е. или $f(\beta)=0$, или же $\overline{f}(\beta)=0$. В первом случае теорема доказана. Если же имеет место второй случай, т. е.

$$\overline{a}_0\beta^n+\overline{a}_1\beta^{n-1}+\ldots+\overline{a}_n=0,$$

то, заменяя все входящие сюда комплексные числа им сопряженными (что, как мы знаем, не нарушает равенства), мы получим:

$$f(\overline{\beta})=a_0\overline{\beta}^n+\alpha_1\overline{\beta}^{n-1}+\ldots+a_n=0,$$

т. е. $f(x)$ имеет своим корнем комплексное число $\overline{\beta}$. Доказательство основной теоремы закончено.

ГЛАВА ДВЕНАДЦАТАЯ

МНОГОЧЛЕНЫ С РАЦИОНАЛЬНЫМИ КОЭФФИЦИЕНТАМИ

§ 56*. Приводимость многочленов над полем рациональных чисел

Третьим числовым полем, которое наряду с полями действительных и комплексных чисел представляет для нас особый интерес, является поле рациональных чисел; обозначим его через R. Оно является самым малым среди числовых полей: как доказано в § 43, поле R содержится целиком во всяком числовом поле. Мы будем интересоваться сейчас вопросом о приводимости многочленов над полем рациональных чисел, а в следующем параграфе — вопросом о рациональных (целых и дробных) корнях многочленов с рациональными коэффициентами. Еще раз подчеркнем, что это два разных вопроса: многочлен

$$x^4+2x^2+1=(x^2+1)^2$$

приводим над полем рациональных чисел, хотя не имеет ни одного рационального корня.

Что можно сказать о приводимости многочленов над полем R? Заметим, прежде всего, что если дан многочлен $f(x)$, коэффициенты которого рациональны, но не все целые, то, приводя коэффициенты к общему знаменателю и умножая $f(x)$ на этот знаменатель, равный, например, k, мы получим многочлен $kf(x)$, все коэффициенты которого будут уже целыми числами. Очевидно, что многочлены $f(x)$ и $kf(x)$ имеют одинаковые корни; с другой стороны, они одновременно будут приводимыми или неприводимыми над полем R.

Мы, однако, пока не получили права ограничиться в дальнейшем рассмотрением многочленов с целыми коэффициентами. В самом деле, пусть целочисленный многочлен $g(x)$ (т. е. многочлен с целыми коэффициентами) приводим над полем рациональных чисел, т. е. разложим на множители меньшей степени с рациональными (вообще говоря, дробными) коэффициентами. Следует ли отсюда разложимость $g(x)$ на множители с целыми коэффициентами? Иными словами, не может ли многочлен с целыми коэффициентами, приводимый над полем рациональных чисел, оказаться неприводимым над кольцом целых чисел?

Ответ на эти вопросы может быть получен при помощи рассмотрений, аналогичных проведенным в § 51. Назовем многочлен $f(x)$ с целыми коэффициентами *примитивным*, если его коэффициенты в совокупности взаимно просты, т. е. не имеют общих делителей, отличных от 1 и —1. Если дан произвольный многочлен $\varphi(x)$ с рациональными коэффициентами, то его можно, притом однозначным образом, представить в виде произведения несократимой дроби на некоторый примитивный многочлен:

$$\varphi(x)=\frac{a}{b}f(x); \tag{1}$$

для этого нужно вынести за скобки общий знаменатель всех коэффициентов многочлена $\varphi(x)$, а затем и общие множители из числителей этих коэффициентов; заметим, что степень $f(x)$ равна степени $\varphi(x)$. Однозначность (с точностью до знака) представления (1) доказывается следующим образом. Пусть

$$\varphi(x)=\frac{a}{b}f(x)=\frac{c}{d}g(x),$$

где $g(x)$ — снова примитивный многочлен. Тогда

$$adf(x)=bcg(x).$$

Таким образом, ad и bc получены вынесением всех общих множителей из коэффициентов одного и того же целочисленного многочлена, а поэтому могут отличаться друг от друга лишь знаком. Отсюда следует, что и примитивные многочлены $f(x)$ и $g(x)$ также могут отличаться друг от друга лишь знаком.

Для целочисленных примитивных многочленов остается справедливой лемма Гаусса:

Произведение двух целочисленных примитивных многочленов само есть примитивный многочлен.

В самом деле, пусть даны примитивные целочисленные многочлены

$$f(x)=a_0x^k+a_1x^{k-1}+\ldots+a_ix^{k-i}+\ldots+a_k,$$
$$g(x)=b_0x^l+b_1x^{l-1}+\ldots+b_jx^{l-j}+\ldots+b_l$$

и пусть

$$f(x)g(x)=c_0x^{k+l}+c_1x^{k+l-1}+\ldots+c_{i+j}x^{(k+l)-(i+j)}+\ldots+c_{k+l}.$$

Если это произведение не примитивно, то существует такое простое число p, которое служит общим делителем для всех коэффициентов $c_0, c_1, \ldots, c_{k+l}$. Так как все коэффициенты примитивного многочлена $f(x)$ не могут делиться на p, то пусть коэффициент a_i будет первым, на p не делящимся; аналогично через b_j мы обозначим первый коэффициент многочлена $g(x)$, не делящийся на p. Перемножая

почленно $f(x)$ и $g(x)$ и собирая члены, содержащие $x^{(k+l)-(i+j)}$, мы получим:

$$c_{i+j}=a_ib_j+a_{i-1}b_{j+1}+a_{i-2}b_{j+2}+\ldots+a_{i+1}b_{j-1}+a_{i+2}b_{j-2}+\ldots$$

Левая часть этого равенства делится на p. На него заведомо делятся также все слагаемые правой части, кроме первого; действительно, ввиду условий, наложенных на выбор i и j, все коэффициенты $a_{i-1}, a_{i-2}, \ldots$, а также $b_{j-1}, b_{j-2}, \ldots$, делятся на p. Отсюда следует, что произведение a_ib_j также делится на p, а поэтому, ввиду простоты числа p, на p должен делиться хотя бы один из коэффициентов a_i, b_j, что, однако, не имеет места. Этим заканчивается доказательство леммы.

Переходим к ответу на поставленные выше вопросы. Пусть многочлен $g(x)$ степени n с целыми коэффициентами приводим над полем рациональных чисел:

$$g(x)=\varphi_1(x)\varphi_2(x),$$

где $\varphi_1(x)$ и $\varphi_2(x)$ — многочлены с рациональными коэффициентами и их степени меньше n. Тогда

$$\varphi_i(x)=\frac{a_i}{b_i}f_i(x), \quad i=1, 2,$$

где $\frac{a_i}{b_i}$ — несократимая дробь, $f_i(x)$ — примитивный многочлен. Отсюда

$$g(x)=\frac{a_1a_2}{b_1b_2}[f_1(x)f_2(x)].$$

Левая часть этого равенства является целочисленным многочленом, поэтому знаменатель b_1b_2 в правой части должен сократиться. Однако многочлен, стоящий в квадратных скобках, будет, по лемме Гаусса, примитивным, поэтому всякий простой множитель из b_1b_2 может сократиться лишь с некоторым простым множителем из a_1a_2, а так как a_i и b_i взаимно просты, $i=1, 2$, то число a_2 должно нацело делиться на b_1, a_1 — на b_2;

$$a_2=b_1a_2', \quad a_1=b_2a_1'.$$

Отсюда

$$g(x)=a_1'a_2'f_1(x)f_2(x).$$

Присоединив коэффициент $a_1'a_2'$ к любому из множителей $f_1(x)$, $f_2(x)$, мы получим разложение многочлена $g(x)$ на множители меньшей степени с целыми коэффициентами. Этим доказана следующая теорема:

Многочлен с целыми коэффициентами, неприводимый над кольцом целых чисел, будет неприводимым и над полем рациональных чисел.

Теперь мы получили, наконец, право ограничиваться в вопросах, относящихся к приводимости многочленов над полем рациональных чисел, рассмотрением разложений целочисленных многочленов на множители, все коэффициенты которых также целые.

Мы знаем, что над полем комплексных чисел приводим всякий многочлен, степень которого больше единицы, а над полем действительных чисел — всякий многочлен (с действительными коэффициентами), степень которого больше двух. Совсем иное положение в случае поля рациональных чисел: *для любого n можно указать многочлен n-й степени с рациональными (даже целыми) коэффициентами, неприводимый над полем рациональных чисел.* Доказательство этого утверждения основано на следующем достаточном признаке неприводимости многочлена над полем R, называемом критерием Эйзенштейна:

Пусть дан многочлен

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

с целыми коэффициентами. Если хотя бы одним способом можно подобрать простое число p, удовлетворяющее следующим требованиям:

1) *старший коэффициент a_0 не делится на p,*

2) *все остальные коэффициенты делятся на p,*

3) *свободный член, делясь на p, не делится на p^2,*

то многочлен $f(x)$ неприводим над полем рациональных чисел.

В самом деле, если многочлен $f(x)$ приводим над полем R, то он разлагается на два множителя меньшей степени с целыми коэффициентами:

$$f(x) = (b_0x^k + b_1x^{k-1} + \ldots + b_k)(c_0x^l + c_1x^{l-1} + \ldots + c_l),$$

где $k < n$, $l < n$, $k + l = n$. Отсюда, сравнивая коэффициенты в обеих частях этого равенства, получаем:

$$\left.\begin{aligned} a_n &= b_kc_l, \\ a_{n-1} &= b_kc_{l-1} + b_{k-1}c_l, \\ a_{n-2} &= b_kc_{l-2} + b_{k-1}c_{l-1} + b_{k-2}c_l, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 &= b_0c_0. \end{aligned}\right\} \qquad (2)$$

Из первого из равенств (2) следует, так как a_n делится на p, а число p простое, что один из множителей b_k, c_l должен делиться на p. Они оба не могут одновременно делиться на p, так как a_n, по условию, не делится на p^2. Пусть, например, b_k делится на p и поэтому c_l взаимно просто с p. Переходим теперь ко второму из равенств (2). Его левая часть, а также первое слагаемое правой части делятся на p, поэтому на p делится и произведение $b_{k-1}c_l$;

так как, однако, c_l на p не делится, то на p будет делиться b_{k-1}. Подобным же образом из третьего равенства (2) мы получим, что b_{k-2} делится на p, и т. д. Наконец, из $(k+1)$-го равенства будет получено, что на p делится b_0; но тогда из последнего из равенств (2) вытекает, что на p делится a_0, что противоречит предположению.

Весьма легко для любого n написать целочисленные многочлены n-й степени, удовлетворяющие условиям критерия Эйзенштейна и, следовательно, неприводимые над полем рациональных чисел. Таков, например, многочлен x^n+2; к нему применим критерий Эйзенштейна при $p=2$.

Критерий Эйзенштейна является лишь достаточным условием неприводимости над полем R, но отнюдь не необходимым: если для данного многочлена $f(x)$ нельзя подобрать такого простого числа p, чтобы выполнялись условия критерия Эйзенштейна, то он может быть приводимым, как x^2-5x+6, но может быть и неприводимым, как x^2+1. Существует, помимо критерия Эйзенштейна, много других достаточных критериев неприводимости многочленов над полем R, впрочем менее значительных. Существует также метод, принадлежащий Кронекеру и позволяющий о любом многочлене с целыми коэффициентами решить, приводим ли он над полем R или нет. Этот метод, однако, очень громоздок и практически почти неприменим.

Пример. Рассмотрим многочлен

$$f_p(x)=\frac{x^p-1}{x-1}=x^{p-1}+x^{p-2}+\ldots+x+1,$$

где p — простое число. Корнями этого многочлена служат корни p-й степени из единицы, отличные от самой единицы; так как эти корни вместе с 1 делят единичный круг комплексной плоскости на p равных частей, то многочлен $f_p(x)$ называется *многочленом деления круга*.

К этому многочлену не может быть непосредственно применен критерий Эйзенштейна. Совершим, однако, замену неизвестного, положив $x=y+1$. Мы получим:

$$g(y)=f_p(y+1)=\frac{(y+1)^p-1}{(y+1)-1}=$$

$$=\frac{1}{y}\left[y^p+py^{p-1}+\frac{p(p-1)}{2!}y^{p-2}+\ldots+py\right]=$$

$$=y^{p-1}+py^{p-2}+\frac{p(p-1)}{2!}y^{p-3}+\ldots+p.$$

Коэффициенты многочлена $g(y)$ являются биномиальными коэффициентами и поэтому все, кроме старшего, делятся на p, причем свободный член не делится на p^2. Таким образом, согласно критерию Эйзенштейна многочлен $g(y)$ неприводим над полем R. Отсюда следует *неприводимость над полем R многочлена деления круга* $f_p(x)$. В самом деле, если

$$f_p(x)=\varphi(x)\psi(x),$$

или

$$g(y)=\varphi(y+1)\psi(y+1).$$

§ 57*. Рациональные корни целочисленных многочленов

Выше было указано, что вопрос о разложении данного многочлена над полем рациональных чисел на неприводимые множители ни имеет практически сколько-нибудь удовлетворительного решения. Однако частный случай этого вопроса, относящийся к выделению линейных множителей многочлена с рациональными коэффициентами, т. е. к разысканию его рациональных корней, уже весьма прост и решается без больших вычислений. Само собой разумеется, что вопрос о разыскании рациональных корней многочленов с рациональными коэффициентами ни в какой мере не исчерпывает общего вопроса о действительных корнях этих многочленов, т. е. методы и результаты, изложенные в девятой главе, сохраняют полностью свое значение и для многочленов с рациональными коэффициентами.

Приступая к вопросу о разыскании рациональных корней многочленов с рациональными коэффициентами, отметим, что, как было указано в предшествующем параграфе, можно ограничиться рассмотрением лишь многочленов с целыми коэффициентами; мы будем при этом рассматривать отдельно случай целых и случай дробных корней.

Если целое число α служит корнем многочлена $f(x)$ с целыми коэффициентами, то α будет делителем свободного члена этого многочлена.

В самом деле, пусть

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n.$$

Разделим $f(x)$ на $x-\alpha$:

$$f(x)=(x-\alpha)(b_0x^{n-1}+b_1x^{n-2}+\ldots+b_{n-1}).$$

Выполняя деление методом Горнера, изложенным в § 22, мы получим, что *все коэффициенты частного, в том числе и b_{n-1}, являются целыми числами*, а так как

$$a_n=-\alpha b_{n-1}=\alpha(-b_{n-1}),$$

то наше утверждение доказано [1]).

Таким образом, если целочисленный многочлен $f(x)$ обладает целыми корнями, то они будут найдены среди делителей свободного члена. Необходимо, следовательно, испытать всевозможные делители свободного члена, как положительные, так и отрицательные; если ни один из них не является корнем многочлена, то целых корней наш многочлен вообще не имеет.

---

[1]) Было бы ошибкой доказывать эту теорему ссылкой на то, что свободный член a_n является (с точностью до знака) произведением всех корней многочлена $f(x)$: среди этих корней могут встретиться и дробные, и иррациональные, и комплексные, и поэтому заранее нельзя утверждать, что произведение всех этих корней, кроме α, будет целым.

Испытание всех делителей свободного члена может оказаться весьма громоздким, если даже значения многочлена будут вычисляться методом Горнера, а не непосредственной подстановкой каждого из делителей вместо неизвестного. Следующие замечания позволяют несколько упростить эти вычисления. Прежде всего, так как 1 и -1 всегда служат делителями свободного члена, вычисляем $f(1)$ и $f(-1)$, что не представляет затруднений. Если, далее, целое число α является корнем для $f(x)$:

$$f(x)=(x-\alpha)\,q(x),$$

то, как указано выше, все коэффициенты частного $q(x)$ будут целыми числами, и поэтому частные

$$\frac{f(1)}{\alpha-1}=-q(1),\quad \frac{f(-1)}{\alpha+1}=-q(-1)$$

должны быть целыми числами. Таким образом, *подлежат испытанию лишь те делители* α *свободного члена* (*из числа отличных от* 1 *и* -1), *для которых каждое из частных* $\frac{f(1)}{\alpha-1}$, $\frac{f(-1)}{\alpha+1}$ *является целым числом.*

Примеры. 1. Найти целые корни многочлена

$$f(x)=x^3-2x^2-x-6.$$

Делителями свободного члена служат числа ±1, ±2, ±3, ±6. Так как $f(1)=-8$, $f(-1)=-8$, то 1 и -1 не являются корнями. Далее, числа

$$\frac{-8}{2+1},\ \frac{-8}{-2-1},\ \frac{-8}{6-1},\ \frac{-8}{-6-1}$$

являются дробными, и поэтому делители 2, -2, 6, -6 должны быть отброшены, в то время как числа

$$\frac{-8}{3-1},\ \frac{-8}{3+1},\ \frac{-8}{-3-1},\ \frac{-8}{-3+1}$$

— целые, и поэтому делители 3 и -3 еще подлежат испытанию. Применим метод Горнера:

$$\begin{array}{r|rrrr} & 1 & -2 & -1 & -6 \\ \hline -3 & 1 & -5 & 14 & -48 \end{array},$$

т. е. $f(-3)=-48$, и поэтому -3 не является корнем для $f(x)$. Наконец

$$\begin{array}{r|rrrr} & 1 & -2 & -1 & -6 \\ \hline 3 & 1 & 1 & 2 & 0 \end{array},$$

т. е. $f(3)=0$: число 3 служит корнем для $f(x)$. Одновременно мы нашли коэффициенты частного от деления $f(x)$ на $x-3$;

$$f(x)=(x-3)(x^2+x+2).$$

Легко видеть, что частное x^2+x+2 не имеет числа 3 своим корнем, т. е. это число не является кратным корнем для $f(x)$.

2. Найти целые корни многочлена

$$f(x)=3x^4+x^3-5x^2-2x+2.$$

Здесь делителями свободного члена будут ± 1 и ± 2. Далее, $f(1)=-1$, $f(-1)=1$, т. е. 1 и -1 не служат корнями. Наконец, так как числа

$$\frac{1}{2+1} \text{ и } \frac{-1}{-2-1}$$

дробные, то 2 и -2 также не будут корнями и поэтому многочлен $f(x)$ вообще не имеет целых корней.

Переходим к вопросу о дробных корнях.

Если целочисленный многочлен, старший коэффициент которого равен единице, имеет рациональный корень, то этот корень будет целым числом.

Пусть, в самом деле, многочлен

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_n$$

с целыми коэффициентами имеет корнем несократимую дробь $\frac{b}{c}$, т. е.

$$\frac{b^n}{c^n}+a_1\frac{b^{n-1}}{c^{n-1}}+a_2\frac{b^{n-2}}{c^{n-2}}+\ldots+a_n=0.$$

Отсюда

$$\frac{b^n}{c}=-a_1b^{n-1}-a_2b^{n-2}c-\ldots-a_nc^{n-1},$$

т. е. несократимая дробь равна целому числу, что невозможно.

Для получения всех рациональных (дробных и целых) корней целочисленного многочлена

$$f(x)=a_0x^n+a_1x^{n-1}+a_2x^{n-2}+\ldots+a_{n-1}x+a_n$$

нужно найти все целые корни многочлена

$$\varphi(y)=y^n+a_1y^{n-1}+a_0a_2y^{n-2}+\ldots+a_0^{n-2}a_{n-1}y+a_0^{n-1}a_n$$

и разделить их на a_0.

В самом деле, умножим $f(x)$ на a_0^{n-1}, а затем совершим замену неизвестного, положив $y=a_0x$. Очевидно, что

$$\varphi(y)=\varphi(a_0x)=a_0^{n-1}f(x).$$

Отсюда следует, что корни многочлена $f(x)$ равны корням многочлена $\varphi(y)$, разделенным на a_0. В частности, рациональным корням $f(x)$ будут соответствовать рациональные же корни $\varphi(y)$; так как, однако, старший коэффициент $\varphi(y)$ равен единице, то эти корни могут быть лишь целыми, и мы уже имеем метод для их разыскания.

Пример. Найти рациональные корни многочлена

$$f(x)=3x^4+5x^3+x^2+5x-2.$$

Умножая $f(x)$ на 3^3 и полагая $y=3x$, получим:

$$\varphi(y)=y^4+5y^3+3y^2+45y-54.$$

Ищем целые корни многочлена $\varphi(y)$.

Найдем $\varphi(1)$ методом Горнера:

$$1\left|\begin{array}{ccccc} 1 & 5 & 3 & 45 & -54 \\ \hline 1 & 6 & 9 & 54 & 0 \end{array}\right. .$$

Таким образом, $\varphi(1)=0$, т. е. 1 является корнем для $\varphi(y)$, причем

$$\varphi(y)=(y-1)\,q(y),$$

где

$$q(y)=y^3+6y^2+9y+54.$$

Найдем целые корни многочлена $q(y)$. Делителями свободного члена служат числа ±1, ±2, ±3, ±6, ±9, ±18, ±27, ±54. Здесь

$$q(1)=70, \quad q(-1)=50.$$

Вычисляя $\frac{q(1)}{\alpha-1}$ и $\frac{q(-1)}{\alpha+1}$ для каждого делителя α, мы обнаружим, что должны быть отброшены все делители, кроме $\alpha=-6$. Испытаем этот делитель:

$$-6\left|\begin{array}{cccc} 1 & 6 & 9 & 54 \\ \hline 1 & 0 & 9 & 0 \end{array}\right. .$$

Таким образом, $q(-6)=0$, т. е. -6 служит корнем для $q(y)$ и поэтому для $\varphi(y)$.

Многочлен $\varphi(y)$ имеет, следовательно, целые корни 1 и -6. Рациональными корнями многочлена $f(x)$ будут, таким образом, числа $\frac{1}{3}$ и -2 и только они.

Следует еще раз подчеркнуть, что изложенные выше методы применимы **только** к многочленам с целыми коэффициентами и **только** для разыскания их рациональных корней.

§ 58*. Алгебраические числа

Всякий многочлен n-й степени с рациональными коэффициентами имеет в поле комплексных чисел n корней, некоторые из которых (или даже все) могут лежать вне поля рациональных чисел. Однако не всякое комплексное или действительное число служит корнем некоторого многочлена с рациональными коэффициентами. Те комплексные (в частности, действительные) числа, которые являются корнями таких многочленов, называются *алгебраическими* числами в противоположность числам *трансцендентным*. К числу алгебраических чисел принадлежат все рациональные числа, как корни многочленов первой степени с рациональными коэффициентами, а также всякий радикал вида $\sqrt[n]{a}$ с рациональным подкоренным числом a, как корень двучлена x^n-a. С другой стороны, в больших курсах математического анализа доказывается трансцендентность числа e — основания системы натуральных логарифмов, а также известного из элементарной геометрии числа π.

Если число α алгебраическое, то оно будет даже корнем некоторого многочлена с целыми коэффициентами и поэтому корнем

одного из неприводимых делителей этого многочлена также с целыми коэффициентами. *Тот неприводимый целочисленный многочлен, корнем которого является α, определен однозначно с точностью до постоянного множителя, т. е. вполне однозначно, если потребовать, чтобы коэффициенты этого многочлена были в совокупности взаимно просты* (т. е. чтобы многочлен был примитивным). В самом деле, если α служит корнем двух неприводимых многочленов $f(x)$ и $g(x)$, то наибольший общий делитель этих многочленов будет отличен от единицы, а потому эти многочлены, ввиду их неприводимости, могут отличаться друг от друга лишь множителем нулевой степени.

Алгебраические числа, являющиеся корнями одного и того же неприводимого (над полем R) многочлена, называются *сопряженными* между собой[1]). Все множество алгебраических чисел распадается, следовательно, на непересекающиеся конечные классы сопряженных между собой чисел. Всякое рациональное число как корень многочлена первой степени не имеет сопряженных чисел, отличных от самого себя, и это свойство является для рациональных чисел характерным: всякое алгебраическое число, не являющееся рациональным, будет корнем неприводимого многочлена, степень которого больше единицы, и поэтому для него существуют сопряженные, отличные от него самого.

Множество всех алгебраических чисел является подполем поля комплексных чисел. Иными словами, *сумма, разность, произведение и частное алгебраических чисел сами будут алгебраическими числами.*

Пусть, в самом деле, даны алгебраические числа α и β. Обозначим через $\alpha_1=\alpha, \alpha_2, \ldots, \alpha_n$ все числа, сопряженные с α, через $\beta_1=\beta, \beta_2, \ldots, \beta_s$ — числа, сопряженные с β, через $f(x)$ и $g(x)$ — неприводимые многочлены с рациональными коэффициентами, имеющие своими корнями соответственно α и β. Напишем многочлен, корнями которого служат всевозможные суммы $\alpha_i+\beta_j$; это будет

$$\varphi(x)=\prod_{i=1}^{n}\prod_{j=1}^{s}[x-(\alpha_i+\beta_j)].$$

Коэффициенты этого многочлена не будут, очевидно, меняться при перестановках всех α_i между собой, а также всех β_j между собой. Они являются, следовательно, на основании теоремы о многочленах, симметричных по двум системам неизвестных (см. конец § 53), многочленами от коэффициентов многочленов $f(x)$ и $g(x)$. Иными словами, коэффициенты многочлена $\varphi(x)$ оказываются рациональными числами, и поэтому число $\alpha+\beta=\alpha_1+\beta_1$, являющееся одним из его корней, будет алгебраическим.

[1]) Не следует смешивать этого понятия с сопряженностью комплексных чисел.

Таким же образом при помощи многочленов

$$\psi(x)=\prod_{i=1}^{n}\prod_{j=1}^{s}[x-(\alpha_i-\beta_j)]$$

и

$$\chi(x)=\prod_{i=1}^{n}\prod_{j=1}^{s}(x-\alpha_i\beta_j)$$

доказывается алгебраичность чисел $\alpha-\beta$ и $\alpha\beta$.

Для доказательства алгебраичности частного достаточно показать, что если число α — алгебраическое и отличное от нуля, то α^{-1} также будет алгебраическим числом. Пусть α служит корнем многочлена

$$f(x)=a_0x^n+a_1x^{n-1}+\ \dots\ +a_{n-1}x+a_n$$

с рациональными коэффициентами. Тогда, очевидно, многочлен

$$g(x)=a_nx^n+a_{n-1}x^{n-1}+\ \dots\ +a_1x+a_0,$$

также с рациональными коэффициентами, будет иметь своим корнем число α^{-1}, что и требовалось доказать.

Из доказанной сейчас теоремы вытекает, что любая сумма рационального числа и радикала, например $1+\sqrt[3]{2}$, а также любая сумма радикалов, например $\sqrt{3}+\sqrt[7]{5}$, будут алгебраическими числами. Мы пока не можем, однако, утверждать алгебраичность чисел, записываемых в виде «двухэтажных» радикалов, например числа $\sqrt{1+\sqrt{2}}$. Это будет вытекать лишь из следующей теоремы:

Если число ω служит корнем многочлена

$$\varphi(x)=x^n+\alpha x^{n-1}+\beta x^{n-2}+\ \dots\ +\lambda x+\mu,$$

коэффициенты которого — алгебраические числа, то ω также будет алгебраическим числом.

Пусть α_i, β_j, ..., λ_s, μ_t пробегают числа, сопряженные соответственно с α, β, ..., λ, μ, причем $\alpha_1=\alpha$, $\beta_1=\beta$, ..., $\lambda_1=\lambda$, $\mu_1=\mu$. Рассмотрим всевозможные многочлены вида

$$\varphi_{i,\,j,\,\dots,\,s,\,t}(x)=x^n+\alpha_ix^{n-1}+\beta_jx^{n-2}+\ \dots\ +\lambda_sx+\mu_t,$$

так что $\varphi_{1,\,1,\,\dots,\,1,\,1}(x)=\varphi(x)$, и возьмем произведение всех этих многочленов

$$F(x)=\prod_{i,\,j,\,\dots,\,s,\,t}\varphi_{i,\,j,\,\dots,\,s,\,t}(x).$$

Коэффициенты многочлена $F(x)$ симметричны, очевидно, по каждой из систем α_i, β_j, ..., λ_s, μ_t, а поэтому (снова по теореме из § 53) они суть многочлены от коэффициентов тех неприводимых многочленов (с рациональными коэффициентами), корнями которых

служат соответственно α, β, ..., λ, μ, т. е. сами суть рациональные числа. Число ω, являясь корнем для $\varphi(x)$, будет, следовательно, корнем многочлена $F(x)$ с рациональными коэффициентами, т. е. будет алгебраическим числом.

Применим эту теорему к числу $\omega=\sqrt{1+\sqrt{2}}$. Число $\alpha=1+\sqrt{2}$ алгебраично по предыдущей теореме и поэтому число ω является корнем многочлена $x^2-\alpha$ с алгебраическими коэффициентами, т. е. само алгебраично. Вообще, применяя несколько раз обе доказанные сейчас теоремы, читатель без труда придет к следующему результату:

Всякое число, записываемое в радикалах над полем рациональных чисел (т. е. выражающееся через сколь угодно сложную комбинацию радикалов, в общем случае «многоэтажных»), будет алгебраическим числом.

Алгебраические числа, записываемые в радикалах, составляют, очевидно, поле. Следует помнить, однако, что это поле, как вытекает из замечания, сделанного (без доказательства) в конце § 38, будет лишь частью поля всех алгебраических чисел.

Выше была отмечена трансцендентность двух чисел: e и π. На самом деле, однако, трансцендентных чисел бесконечно много. Больше того, используя понятия и методы, относящиеся к теории множеств, мы покажем, что трансцендентных чисел, так сказать, даже больше, чем чисел алгебраических; точный смысл этого выражения станет ясен ниже.

Бесконечное множество M называется *счетным*, если оно может быть поставлено во взаимно однозначное соответствие с множеством натуральных чисел, т. е. если его элементы могут быть пронумерованы при помощи всех натуральных чисел, и *несчетным* — в противоположном случае.

Лемма 1. *Всякое бесконечное множество M содержит счетное подмножество.*

В самом деле, возьмем в M произвольный элемент a_1. Выберем затем элемент a_2, отличный от a_1. Вообще, пусть в M уже выбрано n различных элементов $a_1, a_2, \ldots, a_n$. Так как множество M, будучи бесконечным, не может исчерпываться этими элементами, то можно указать отличный от них элемент a_{n+1}. Продолжая этот процесс, мы найдем в M бесконечное подмножество, составленное из элементов

$$a_1, a_2, \ldots, a_n, \ldots;$$

счетность этого подмножества очевидна.

Лемма 2. *Всякое бесконечное подмножество B счетного множества A само счетно.*

Множество A, ввиду его счетности, можно записать в виде

$$a_1, a_2, \ldots, a_n, \ldots \quad (1)$$

Пусть a_{k_1} будет первый элемент последовательности (1), принадлежащий к B, a_{k_2} — второй элемент с этим же свойством и т. д. Полагая $a_{k_n}=b_n$, $n=1, 2, \ldots$, мы получаем, что элементы подмножества B составляют последовательность

$$b_1, b_2, \ldots, b_n, \ldots,$$

т. е. это подмножество счетное.

Лемма 3. *Объединение счетного множества конечных множеств, попарно не имеющих общих элементов, есть счетное множество.*

Пусть, в самом деле, даны конечные множества

$$A_1, A_2, \ldots, A_n, \ldots$$

и пусть их объединение будет B. Мы пронумеруем, очевидно, все элементы множества B, если произвольным образом пронумеруем элементы конечного множества A_1, затем продолжим эту нумерацию, перейдя к элементам множества A_2, и т. д.

Лемма 4. *Объединение двух счетных множеств, не имеющих общих элементов, есть счетное множество.*

Пусть даны счетные множества A с элементами

$$a_1, a_2, \ldots, a_n, \ldots$$

и B с элементами

$$b_1, b_2, \ldots, b_n, \ldots$$

и пусть объединение этих множеств будет C. Если мы положим

$$a_n = c_{2n-1}, \quad b_n = c_{2n}, \quad n = 1, 2, \ldots,$$

то все элементы множества C будут представлены в виде последовательности

$$c_1, c_2, \ldots, c_{2n-1}, c_{2n}, \ldots,$$

что и доказывает счетность этого множества.

Докажем теперь следующую теорему:

Множество всех алгебраических чисел счетно.

Докажем предварительно *счетность множества всех многочленов от одного неизвестного с целыми коэффициентами.* Если

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

— такой многочлен, притом отличный от нуля, то назовем *высотой* этого многочлена натуральное число

$$h_f = n + |a_0| + |a_1| + \ldots + |a_{n-1}| + |a_n|.$$

Очевидно, что существует лишь конечное число целочисленных многочленов с данной высотой h; обозначим это множество через M_h. Кроме того, через M_0 обозначим множество, состоящее из одного нуля. Множество всех целочисленных многочленов будет объединением счетного множества конечных множеств $M_0, M_1, M_2, \ldots, M_h, \ldots$, т. е., по лемме 3, оно счетно.

Отсюда, по лемме 2, вытекает, что *множество всех целочисленных примитивных неприводимых многочленов также счетно.* Мы знаем, вместе с тем, что всякое алгебраическое число является корнем одного и только одного целочисленного примитивного неприводимого многочлена. Собирая, следовательно, корни всех таких многочленов, т. е. беря объединение счетного множества конечных множеств, мы получим множество всех алгебраических чисел; это множество будет, таким образом, ввиду леммы 3, счетным.

Докажем, наконец, теорему:

Множество всех трансцендентных чисел несчетно.

Рассмотрим сначала множество F всех действительных чисел x, расположенных между нулем и единицей, $0 < x < 1$, и докажем, что *это множество несчетно.* Известно, что каждое из указанных чисел x можно записать в виде правильной бесконечной десятичной дроби

$$x = 0, \alpha_1\alpha_2 \ldots \alpha_n \ldots$$

и что эта запись однозначна, если не допускать дробей, у которых для всех n, начиная с некоторого $n = N$, все $\alpha_n = 9$; обратно, всякая дробь указанного

вида равна некоторому числу x из множества F. Предположим теперь, что множество F счетно, т. е. что все числа x можно записать в виде последовательности

$$x_1, x_2, \ldots, x_k, \ldots \tag{2}$$

Пусть

$$x_k = 0, \alpha_{k_1}\alpha_{k_2} \ldots \alpha_{kn} \ldots$$

будет запись числа x_k в виде бесконечной десятичной дроби. Напишем теперь бесконечную десятичную дробь

$$0, \beta_1\beta_2 \ldots \beta_n \ldots, \tag{3}$$

полагая β_1 отличным от первого десятичного знака дроби x_1, т. е. $\beta_1 \neq \alpha_{11}$, β_2 — отличным от второго десятичного знака дроби x_2, т. е. $\beta_2 \neq \alpha_{22}$, и, вообще, $\beta_n \neq \alpha_{nn}$. Положим, кроме того, что среди цифр β_n бесконечно много отличных от цифры 9. Ясно, что существует дробь (3), удовлетворяющая всем этим требованиям. Она является, следовательно, числом из множества F, но, по самому построению, отлична от всех чисел последовательности (2). Это противоречие доказывает несчетность множества F.

Отсюда следует *несчетность множества всех комплексных чисел:* если бы оно было счетным, то, ввиду леммы 2, оно не могло бы содержать несчетного подмножества F. Несчетность множества всех трансцендентных чисел теперь, ввиду леммы 4, очевидна, так как объединение этого множества со счетным множеством всех алгебраических чисел является множеством всех комплексных чисел, т. е. несчетно.

Доказанные нами две теоремы показывают, ввиду леммы 1, что множество трансцендентных чисел на самом деле является много более богатым элементами, т. е. более «мощным», чем множество алгебраических чисел

ГЛАВА ТРИНАДЦАТАЯ

НОРМАЛЬНАЯ ФОРМА МАТРИЦЫ

§ 59. Эквивалентность λ-матриц

Мы еще раз возвращаемся к вопросам, относящимся к линейной алгебре. Читатель уже при изучении гл. 7 убедился в том, какую важную роль играет понятие подобия матриц. Именно, две квадратные матрицы порядка n подобны тогда и только тогда, если они задают (в разных базах) одно и то же линейное преобразование n-мерного линейного пространства. Мы пока не умеем, однако, отвечать на вопрос, подобны ли две данные конкретные матрицы или нет. С другой стороны, мы пока не умеем находить среди всех матриц, подобных данной матрице A, матрицу, имеющую в том или ином смысле простейший вид, и даже вопрос об условиях, при которых матрица A подобна диагональной матрице, был рассмотрен в § 33 лишь в одном частном случае. Именно эти вопросы будут рассматриваться в настоящей главе, причем сразу для случая произвольного основного поля P.

Займемся сначала изучением квадратных матриц порядка n, элементами которых служат многочлены произвольных степеней от одного неизвестного λ с коэффициентами из поля P. Такие матрицы называются *многочленными матрицами, полиномиальными матрицами* или, короче, *λ-матрицами*. Примером λ-матрицы служит характеристическая матрица $A-\lambda E$ произвольной квадратной матрицы A с элементами из поля P; на главной диагонали этой матрицы стоят многочлены первой степени, вне главной диагонали — многочлены нулевой степени или нули. Всякая матрица с элементами из поля P — такие матрицы для краткости будем называть *числовыми матрицами* — также будет частным случаем λ-матрицы: ее элементы являются многочленами нулевой степени или нулями.

Пусть дана λ-матрица

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & \dots & a_{1n}(\lambda) \\ \cdot & \cdot & \cdot \\ a_{n1}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

Назовем *элементарными преобразованиями* этой матрицы преобразования следующих четырех типов:

1) умножение любой строки матрицы $A(\lambda)$ на любое число α из поля P, отличное от нуля;

2) умножение любого столбца матрицы $A(\lambda)$ на любое число α из поля P, отличное от нуля;

3) прибавление к любой i-й строке матрицы $A(\lambda)$ любой ее j-й строки, $j \neq i$, притом умноженной на любой многочлен $\varphi(\lambda)$ из кольца $P[\lambda]$;

4) прибавление к любому i-му столбцу матрицы $A(\lambda)$ любого ее j-го столбца, $j \neq i$, притом умноженного на любой многочлен $\varphi(\lambda)$ из кольца $P[\lambda]$.

Легко видеть, что *для каждого из элементарных преобразований λ-матрицы существует обратное преобразование, также являющееся элементарным.* Так, обратным для преобразования 1) будет элементарное преобразование, состоящее в умножении той же строки на число α^{-1}, существующее ввиду условия $\alpha \neq 0$; обратным для преобразования 3) будет преобразование, состоящее в прибавлении к i-й строке j-й строки, умноженной на $-\varphi(\lambda)$.

В матрице $A(\lambda)$ можно при помощи нескольких элементарных преобразований переставить любые две строки или любые два столбца.

Пусть, например, нужно переставить i-ю и j-ю строки матрицы $A(\lambda)$. Это можно сделать при помощи четырех элементарных преобразований, как показывает следующая схема:

$$\begin{pmatrix} i \\ j \end{pmatrix} \to \begin{pmatrix} i+j \\ j \end{pmatrix} \to \begin{pmatrix} i+j \\ -i \end{pmatrix} \to \begin{pmatrix} j \\ -i \end{pmatrix} \to \begin{pmatrix} j \\ i \end{pmatrix}.$$

Здесь последовательно выполнялись такие преобразования: а) к i-й строке прибавлялась j-я; б) из j-й строки вычиталась новая i-я; в) к новой i-й строке прибавлялась новая j-я; г) новая j-я строка умножалась на -1.

Будем говорить, что λ-матрицы $A(\lambda)$ и $B(\lambda)$ *эквивалентны* и записывать это символом $A(\lambda) \sim B(\lambda)$, если от матрицы $A(\lambda)$ можно перейти к матрице $B(\lambda)$ при помощи конечного числа элементарных преобразований. Это отношение эквивалентности является, очевидно, рефлексивным и транзитивным, а также и симметричным ввиду существования для каждого элементарного преобразования обратного элементарного преобразования. Иными словами, *все квадратные λ-матрицы порядка n над полем P распадаются на непересекающиеся классы эквивалентных матриц.*

Нашей ближайшей целью является разыскание среди всех λ-матриц, эквивалентных данной матрице $A(\lambda)$, матрицы по возможности простого вида. Для этого введем следующее понятие. *Канонической λ-матрицей* называется λ-матрица, обладающая следующими тремя свойствами:

а) эта матрица диагональная, т. е. имеет вид

$$\begin{pmatrix} e_1(\lambda) & & & 0 \\ & e_2(\lambda) & & \\ & & \ddots & \\ 0 & & & e_n(\lambda) \end{pmatrix}; \qquad (1)$$

б) всякий многочлен $e_i(\lambda)$, $i=2, 3, \ldots, n$, нацело делится на многочлен $e_{i-1}(\lambda)$;

в) старший коэффициент каждого многочлена $e_i(\lambda)$, $i=1, 2, \ldots, n$, равен единице, если этот многочлен отличен от нуля.

Заметим, что если среди многочленов $e_i(\lambda)$, стоящих на главной диагонали канонической λ-матрицы (1), встречаются равные нулю, то, ввиду свойства б), они непременно занимают на главной диагонали последние места. С другой стороны, если среди многочленов $e_i(\lambda)$ встречаются многочлены нулевой степени, то, по свойству в), все они равны 1 и, по свойству б), занимают на главной диагонали матрицы (1) первые места.

К числу канонических λ-матриц принадлежат, в частности, некоторые числовые матрицы, в том числе матрицы единичная и нулевая.

Всякая λ-матрица эквивалентна некоторой канонической λ-матрице, т. е., иными словами, она приводится элементарными преобразованиями к каноническому виду.

Будем доказывать эту теорему индукцией по порядку n рассматриваемых λ-матриц. Действительно, при $n=1$ будет

$$A(\lambda)=(a(\lambda)).$$

Если $a(\lambda)=0$, то наша матрица уже каноническая. Если же $a(\lambda)\neq 0$, то достаточно разделить многочлен $a(\lambda)$ на его старший коэффициент — это будет элементарное преобразование матрицы — и мы получим каноническую матрицу.

Пусть теорема уже доказана для λ-матриц порядка $n-1$. Рассмотрим произвольную λ-матрицу $A(\lambda)$ порядка n. Если она нулевая, то уже является канонической и доказывать нечего. Будем считать поэтому, что среди элементов матрицы $A(\lambda)$ имеются ненулевые.

Переставляя, если понадобится, строки и столбцы матрицы $A(\lambda)$, можно перевести один из ненулевых элементов в левый верхний угол. Таким образом, среди λ-матриц, эквивалентных матрице $A(\lambda)$, имеются такие, в левом верхнем углу которых стоит ненулевой многочлен. Рассмотрим все такие матрицы. Многочлены, стоящие в левом верхнем углу этих матриц, могут иметь разные степени. Степень многочлена является, однако, натуральным числом, а во всяком непустом множестве натуральных чисел существует наимень-

шее число. Можно найти, следовательно, среди всех λ-матриц, эквивалентных матрице $A(\lambda)$ и имеющих ненулевой элемент в левом верхнем углу, одну из таких, что многочлен, стоящий в ее левом верхнем углу, имеет наименьшую возможную степень. Деля, наконец, первую строку этой матрицы на старший коэффициент указанного многочлена, мы получим такую λ-матрицу, эквивалентную матрице $A(\lambda)$,

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ b_{21}(\lambda) & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1}(\lambda) & b_{n2}(\lambda) & \dots & b_{nn}(\lambda) \end{pmatrix},$$

что $e_1(\lambda) \neq 0$, старший коэффициент этого многочлена равен 1 и никакой комбинацией элементарных преобразований нельзя перейти от полученной матрицы к такой матрице, в левом верхнем углу которой стоял бы ненулевой многочлен меньшей степени.

Докажем, что *все элементы первой строки и первого столбца полученной матрицы нацело делятся на* $e_1(\lambda)$. Пусть, например, для $2 \leqslant j \leqslant n$

$$b_{1j}(\lambda) = e_1(\lambda)\,q(\lambda) + r(\lambda),$$

где степень $r(\lambda)$ меньше степени $e_1(\lambda)$, если $r(\lambda)$ отлично от нуля. Тогда, вычитая из j-го столбца нашей матрицы ее первый столбец, умноженный на $q(\lambda)$, а затем переставляя первый и j-й столбцы, мы придем к такой матрице, эквивалентной матрице $A(\lambda)$, в левом верхнем углу которой стоит многочлен $r(\lambda)$, т. е. многочлен меньшей степени, чем $e_1(\lambda)$, что противоречит выбору этого многочлена. Отсюда следует $r(\lambda) = 0$, что и требовалось доказать.

Вычитая теперь из j-го столбца нашей матрицы ее первый столбец, умноженный на $q(\lambda)$, мы заменим элемент $b_{1j}(\lambda)$ нулем. Делая такие преобразования для $j = 2, 3, \dots, n$, мы заменим нулями все элементы $b_{1j}(\lambda)$. Аналогичным путем заменяются нулями и все элементы $b_{i1}(\lambda)$, $i = 2, 3, \dots, n$. *Мы придем, следовательно, к такой матрице, эквивалентной матрице* $A(\lambda)$, *в левом верхнем углу которой стоит многочлен* $e_1(\lambda)$, *а все остальные элементы первой строки и первого столбца равны нулю,*

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \cdot & \cdot & \cdot & \cdot \\ 0 & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix}. \qquad (2)$$

По индуктивному предположению, матрица $(n-1)$-го порядка, стоящая в правом нижнем углу полученной нами матрицы (2),

элементарными преобразованиями приводится к каноническому виду:

$$\begin{pmatrix} c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \cdot & \cdot & \cdot \\ c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{pmatrix} \sim \begin{pmatrix} e_2(\lambda) & & 0 \\ & \ddots & \\ 0 & & e_n(\lambda) \end{pmatrix}.$$

Совершив эти же преобразования над соответствующими строками и столбцами матрицы (2) — при этом первая строка и первый столбец этой матрицы останутся, очевидно, без изменения, — мы получим, что

$$A(\lambda) \sim \begin{pmatrix} e_1(\lambda) & & & 0 \\ & e_2(\lambda) & & \\ & & \ddots & \\ 0 & & & e_n(\lambda) \end{pmatrix}. \qquad (3)$$

Для доказательства того, что матрица (3) является канонической, остается показать, что $e_2(\lambda)$ нацело делится на $e_1(\lambda)$. Пусть

$$e_2(\lambda) = e_1(\lambda)\, q(\lambda) + r(\lambda),$$

где $r(\lambda) \neq 0$ и степень $r(\lambda)$ меньше степени $e_1(\lambda)$. Прибавляя, однако, ко второму столбцу матрицы (3) ее первый столбец, умноженный на $q(\lambda)$, а затем вычитая из второй строки первую строку, мы заменим элемент $e_2(\lambda)$ элементом $r(\lambda)$. Переставляя, далее, первые две строки и первые два столбца, мы переместим многочлен $r(\lambda)$ в левый верхний угол матрицы, что противоречит, однако, выбору многочлена $e_1(\lambda)$.

Теорема о приведении λ-матрицы к каноническому виду доказана. Эта теорема должна быть дополнена следующей теоремой единственности:

Всякая λ-матрица эквивалентна лишь одной канонической матрице.

В самом деле, пусть дана произвольная λ-матрица $A(\lambda)$ порядка n. Фиксируем некоторое натуральное число k, $1 \leqslant k \leqslant n$, и рассмотрим все миноры k-го порядка матрицы $A(\lambda)$. Вычисляя эти миноры, мы получим конечную систему многочленов от λ; наибольший общий делитель этой системы многочленов, взятый со старшим коэффициентом 1, обозначим через $d_k(\lambda)$.

Мы имеем, следовательно, многочлены

$$d_1(\lambda),\ d_2(\lambda),\ \dots,\ d_n(\lambda), \qquad (4)$$

однозначно определяемые самой матрицей $A(\lambda)$. При этом $d_1(\lambda)$ есть наибольший общий делитель всех элементов матрицы $A(\lambda)$, взятый с коэффициентом 1, а $d_n(\lambda)$ равен определителю матрицы $A(\lambda)$, деленному на его старший коэффициент. Заметим также, что если матрица $A(\lambda)$ имеет ранг r, то

$$d_{r+1}(\lambda) = \dots = d_n(\lambda) = 0,$$

в то время как все остальные многочлены системы (4) отличны от нуля.

Наибольший общий делитель $d_k(\lambda)$ *всех миноров* k*-го порядка* λ*-матрицы* $A(\lambda)$, $k=1, 2, \ldots, n$, *не меняется при выполнении в матрице* $A(\lambda)$ *элементарных преобразований.*

Это утверждение почти очевидно для того случая, когда в матрице $A(\lambda)$ выполняется элементарное преобразование типа 1) или 2). Так, например, если i-я строка матрицы умножается на число α из поля P, $\alpha \neq 0$, то те миноры k-го порядка, через которые i-я строка проходит, будут умножаться на α, все же остальные миноры k-го порядка останутся без изменения. Однако при разыскании наибольшего общего делителя нескольких многочленов любые из этих многочленов можно беспрепятственно умножать на отличные от нуля числа из поля P.

Рассмотрим теперь элементарные преобразования типа 3) или 4). Пусть, например, к i-й строке матрицы $A(\lambda)$ прибавляется ее j-я строка, $j \neq i$, умноженная на многочлен $\varphi(\lambda)$; получающуюся после этого преобразования матрицу обозначим через $\overline{A}(\lambda)$, а наибольший общий делитель всех ее миноров k-го порядка, взятый со старшим коэффициентом 1, — через $\overline{d}_k(\lambda)$. Посмотрим, что происходит при указанном преобразовании с минорами k-го порядка матрицы $A(\lambda)$.

Ясно, что не будут меняться те миноры, через которые i-я строка не проходит. Не меняются и те миноры, через которые проходят как i-я, так и j-я строки, так как определитель не меняется от прибавления к одной его строке кратного другой его строки. Возьмем, наконец, любой из тех миноров k-го порядка, через которые проходит i-я строка, но не проходит j-я; обозначим его через M. Соответствующий минор матрицы $\overline{A}(\lambda)$ можно представить, очевидно, как сумму минора M и умноженного на $\varphi(\lambda)$ минора M' матрицы $A(\lambda)$, получающегося из минора M заменой элементов i-й строки матрицы $A(\lambda)$ соответствующими элементами ее j-й строки. Так как и M, и M' делятся на $d_k(\lambda)$, то и $M+\varphi(\lambda)M'$ будет делиться на $d_k(\lambda)$.

Из сказанного следует, что все миноры k-го порядка матрицы $\overline{A}(\lambda)$ нацело делятся на $d_k(\lambda)$, а поэтому и $\overline{d}_k(\lambda)$ делится на $d_k(\lambda)$. Так как, однако, для рассматриваемого элементарного преобразования существует обратное элементарное преобразование того же типа, то и $d_k(\lambda)$ делится на $\overline{d}_k(\lambda)$. Если же учесть, что старшие коэффициенты обоих этих многочленов равны 1, то $\overline{d}_k(\lambda)=d_k(\lambda)$, что и требовалось доказать.

Таким образом, *всем* λ*-матрицам, эквивалентным матрице* $A(\lambda)$, *соответствует один и тот же набор многочленов* (4). Это относится, в частности, к любой (если их несколько) канонической матрице, эквивалентной $A(\lambda)$. Пусть (3) будет одна из таких матриц.

Вычислим многочлен $d_k(\lambda)$, $k=1, 2, \ldots, n$, пользуясь матрицей (3). Ясно, что минор k-го порядка, стоящий в левом верхнем углу этой матрицы, равен произведению

$$e_1(\lambda)\, e_2(\lambda) \ldots e_k(\lambda). \tag{5}$$

Если, далее, мы берем в матрице (3) минор k-го порядка, стоящий в строках с номерами $i_1, i_2, \ldots, i_k$, где $i_1 < i_2 < \ldots < i_k$, и в столбцах с теми же самыми номерами, то этот минор равен произведению $e_{i_1}(\lambda)\, e_{i_2}(\lambda) \ldots e_{i_k}(\lambda)$, которое делится на (5). Действительно, $1 \leqslant i_1$ и поэтому $e_{i_1}(\lambda)$ делится на $e_1(\lambda)$, $2 \leqslant i_2$, и поэтому $e_{i_2}(\lambda)$ делится на $e_2(\lambda)$ и т. д. Наконец, если в матрице (3) взят минор k-го порядка, через который хотя бы для одного i проходит i-я строка этой матрицы, но не проходит ее i-й столбец, то этот минор содержит нулевую строку и поэтому равен нулю.

Из сказанного следует, что произведение (5) и будет наибольшим общим делителем всех миноров k-го порядка матрицы (3), а поэтому и исходной матрицы $A(\lambda)$,

$$d_k(\lambda) = e_1(\lambda)\, e_2(\lambda) \ldots e_k(\lambda), \qquad k=1, 2, \ldots, n. \tag{6}$$

Теперь легко показать, что *многочлены* $e_k(\lambda)$, $k=1, 2, \ldots, n$, *однозначным образом определяются самой матрицей* $A(\lambda)$. Пусть ранг этой матрицы равен r. Тогда, как мы знаем, $d_r(\lambda) \neq 0$, но $d_{r+1}(\lambda)=0$, а поэтому, ввиду (6), $e_{r+1}(\lambda)=0$. Отсюда, ввиду свойств канонической матрицы, вообще следует, что если ранг r матрицы $A(\lambda)$ меньше n, то

$$e_{r+1}(\lambda) = e_{r+2}(\lambda) = \ldots = e_n(\lambda) = 0. \tag{7}$$

С другой стороны, для $k \leqslant r$ из (6) следует, ввиду $d_{k-1}(\lambda) \neq 0$, что

$$e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)}. \tag{8}$$

Этим заканчивается доказательство единственности канонического вида λ-матрицы. Одновременно мы получили способ непосредственного разыскания многочленов $e_k(\lambda)$, называемых *инвариантными множителями* матрицы $A(\lambda)$.

Пример. Привести к каноническому виду λ-матрицу

$$A(\lambda) = \begin{pmatrix} \lambda^3-\lambda & 2\lambda^2 \\ \lambda^2+5\lambda & 3\lambda \end{pmatrix}.$$

Выполняя цепочку элементарных преобразований, получаем:

$$A(\lambda) \sim \begin{pmatrix} \lambda^3-\lambda & \frac{2}{3}\lambda^2 \\ \lambda^2+5\lambda & \lambda \end{pmatrix} \sim \begin{pmatrix} \frac{1}{3}\lambda^3 - \frac{10}{3}\lambda^2 - \lambda & 0 \\ \lambda^2+5\lambda & \lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} \frac{1}{3}\lambda^3 - \frac{10}{3}\lambda^2 - \lambda & 0 \\ 0 & \lambda \end{pmatrix} \sim \begin{pmatrix} \lambda^3-10\lambda^2-3\lambda & 0 \\ 0 & \lambda \end{pmatrix} \sim \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^3-10\lambda^2-3\lambda \end{pmatrix}.$$

С другой стороны, можно было бы непосредственно вычислить инвариантные множители матрицы $A(\lambda)$. Именно, вычисляя наибольший общий делитель элементов этой матрицы, получаем:

$$d_1(\lambda) = e_1(\lambda) = \lambda.$$

Вычисляя же определитель матрицы $A(\lambda)$ и замечая, что его старший коэффициент равен 1, получаем:

$$d_2(\lambda) = \lambda^4 - 10\lambda^3 - 3\lambda^2,$$

а поэтому

$$e_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = \lambda^3 - 10\lambda^2 - 3\lambda.$$

§ 60. Унимодулярные λ-матрицы. Связь подобия числовых матриц с эквивалентностью их характеристических матриц

Из результатов предшествующего параграфа вытекает один критерий эквивалентности λ-матриц, которому можно придать следующие две почти тождественные формулировки:

Две λ-матрицы тогда и только тогда эквивалентны, если они приводятся к одному и тому же каноническому виду.

Две λ-матрицы тогда и только тогда эквивалентны, если они обладают одинаковыми инвариантными множителями.

Выведем еще один критерий, имеющий уже иной характер.

Мы знаем, что к числу канонических λ-матриц принадлежит единичная матрица E. Назовем λ-матрицу $U(\lambda)$ *унимодулярной*, если она имеет матрицу E своим каноническим видом, т. е. если все ее инвариантные множители равны единице.

λ-матрица $U(\lambda)$ тогда и только тогда унимодулярна, если ее определитель отличен от нуля, но не зависит от λ, т. е. является отличным от нуля числом из основного поля P.

Действительно, если $U(\lambda) \sim E$, то этим двум матрицам соответствует один и тот же многочлен $d_n(\lambda)$. Однако для единичной матрицы $d_n(\lambda) = 1$. Отсюда следует, что определитель матрицы $U(\lambda)$, отличающийся от $d_n(\lambda)$ лишь отличным от нуля числовым множителем, будет отличным от нуля числом из поля P. Обратно, если определитель матрицы $U(\lambda)$ отличен от нуля и не зависит от λ, то для этой матрицы многочлен $d_n(\lambda)$ будет равен 1, а поэтому, по (6) из предыдущего параграфа, все инвариантные множители $e_i(\lambda)$ матрицы $U(\lambda)$, $i = 1, 2, \ldots, n$, равны единице.

Отсюда следует, что *всякая невырожденная числовая матрица является унимодулярной λ-матрицей.* Унимодулярная λ-матрица может иметь, однако, очень сложный вид. Так, λ-матрица

$$\begin{pmatrix} \lambda & \lambda^3 + 5 \\ \lambda^2 - \lambda - 4 & \lambda^4 - \lambda^3 - 4\lambda^2 + 5\lambda - 5 \end{pmatrix}$$

унимодулярна, так как ее определитель равен 20, т. е. отличен от нуля и от λ не зависит.

Из доказанной выше теоремы следует, что *произведение унимодулярных λ-матриц само унимодулярно* — достаточно вспомнить, что при умножении матриц их определители перемножаются.

λ-матрица $U(\lambda)$ тогда и только тогда унимодулярна, если для нее существует обратная матрица, также являющаяся λ-матрицей.

Действительно, если дана невырожденная λ-матрица, то, разыскивая обычным способом обратную матрицу, мы должны будем делить алгебраические дополнения к элементам данной матрицы на определитель этой матрицы, т. е. на некоторый многочлен от λ. Поэтому в общем случае элементы обратной матрицы будут рациональными дробями от λ, а не многочленами от λ, т. е. эта матрица не будет λ-матрицей. Если же дана унимодулярная матрица, то делить алгебраические дополнения придется лишь на отличное от нуля число из поля P, т. е. элементы обратной матрицы будут многочленами от λ и поэтому обратная матрица сама будет λ-матрицей. Обратно, если λ-матрица $U(\lambda)$ обладает обратной λ-матрицей $U^{-1}(\lambda)$, то определители этих обеих матриц являются многочленами от λ, их произведение равно 1, а поэтому оба определителя должны быть многочленами нулевой степени.

Из последнего замечания вытекает такое добавление к доказанной сейчас теореме:

λ-матрица, обратная к унимодулярной λ-матрице, сама унимодулярна.

Понятие унимодулярной матрицы используется в формулировке следующего нового критерия эквивалентности λ-матриц:

Две λ-матрицы $A(\lambda)$ и $B(\lambda)$ порядка n тогда и только тогда эквивалентны, если существуют такие унимодулярные λ-матрицы $U(\lambda)$ и $V(\lambda)$ того же порядка n, что

$$B(\lambda) = U(\lambda)\, A(\lambda)\, V(\lambda). \tag{1}$$

Введем сначала следующее понятие, используемое при доказательстве этого критерия. Назовем *элементарной матрицей* числовую (и, следовательно, λ-) матрицу

$$\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ \cdots & \cdots & \cdots & \alpha & \cdots & \cdots & \cdots \\ & & & & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} (i), \tag{2}$$

отличающуюся от единичной матрицы лишь тем, что на некотором i-м месте главной диагонали, $1 \leqslant i \leqslant n$, стоит произвольное число α

из поля P, отличное от нуля. С другой стороны, назовем также *элементарной матрицей* λ-матрицу

$$\begin{pmatrix} 1 & & & \cdot & & & \\ & \ddots & & \vdots & & & \\ \dots & & 1 & \dots\varphi(\lambda)\dots & & & \\ & & & \ddots & & & \\ & & & 1 & & & \\ & & & \vdots & \ddots & & \\ & & & \vdots & & & 1 \end{pmatrix} (i), \qquad (3)$$

$$(j)$$

отличающуюся от единичной матрицы лишь тем, что на пересечении i-й строки и j-го столбца, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$, причем $i \neq j$, стоит произвольный многочлен $\varphi(\lambda)$ из кольца $P[\lambda]$.

Всякая элементарная матрица унимодулярна. В самом деле, определитель матрицы (2) равен α, но, по условию, $\alpha \neq 0$; определитель же матрицы (3) равен 1.

Выполнение в λ-матрице $A(\lambda)$ любого элементарного преобразования равносильно умножению этой матрицы слева или справа на некоторую элементарную матрицу.

Действительно, читатель без труда проверит справедливость следующих четырех утверждений: 1) умножение матрицы $A(\lambda)$ слева на матрицу (2) равносильно умножению i-й строки матрицы $A(\lambda)$ на число α; 2) умножение матрицы $A(\lambda)$ справа на матрицу (2) равносильно умножению i-го столбца матрицы $A(\lambda)$ на число α; 3) умножение матрицы $A(\lambda)$ слева на матрицу (3) равносильно прибавлению к i-й строке матрицы $A(\lambda)$ ее j-й строки, умноженной на $\varphi(\lambda)$; 4) умножение матрицы $A(\lambda)$ справа на матрицу (3) равносильно прибавлению к j-му столбцу матрицы $A(\lambda)$ ее i-го столбца, умноженного на $\varphi(\lambda)$.

Перейдем теперь к доказательству нашего критерия эквивалентности λ-матриц. Если $A(\lambda) \sim B(\lambda)$, то от $A(\lambda)$ можно перейти к $B(\lambda)$ при помощи конечного числа элементарных преобразований. Заменяя каждое из этих преобразований умножением слева или справа на элементарную матрицу, мы придем к равенству

$$B(\lambda) = U_1(\lambda) \dots U_k(\lambda) A(\lambda) V_1(\lambda) \dots V_l(\lambda), \qquad (4)$$

где все матрицы $U_1(\lambda), \dots, U_k(\lambda), V_1(\lambda), \dots, V_l(\lambda)$ элементарны и, следовательно, унимодулярны. Унимодулярными будут поэтому и матрицы

$$U(\lambda) = U_1(\lambda) \dots U_k(\lambda), \qquad V(\lambda) = V_1(\lambda) \dots V_l(\lambda), \qquad (5)$$

являющиеся произведениями унимодулярных матриц, а равенство (4)

перепишется в виде (1). Заметим, что если, например, $k=0$, т. е. элементарные преобразования совершались лишь над столбцами, то полагаем просто $U(\lambda)=E$.

Проведенная нами часть доказательства позволяет одновременно высказать следующее утверждение:

λ-матрица тогда и только тогда унимодулярна, если она представима в виде произведения элементарных матриц.

В самом деле, мы уже пользовались тем, что произведение элементарных матриц унимодулярно. Обратно, если дана произвольная унимодулярная матрица $W(\lambda)$, то она эквивалентна единичной матрице E. Применяя проведенное выше доказательство вместо матриц $A(\lambda)$ и $B(\lambda)$ к матрицам E и $W(\lambda)$, мы из (4) получим равенство

$$W(\lambda)=U_1(\lambda)\ \dots\ U_k(\lambda)\,V_1(\lambda)\ \dots\ V_l(\lambda),$$

т. е. матрица $W(\lambda)$ оказалась представленной в виде произведения элементарных матриц.

Теперь легко провести и доказательство обратного утверждения нашего критерия. Пусть для матриц $A(\lambda)$ и $B(\lambda)$ существуют такие унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$, что имеет место равенство (1). По доказанному, матрицы $U(\lambda)$ и $V(\lambda)$ можно представить в виде произведений элементарных матриц; пусть это будут представления (5). Равенство (1) перепишется теперь в виде (4) и, заменяя каждое умножение на элементарную матрицу соответствующим элементарным преобразованием, мы получим, наконец, что $A(\lambda)\sim B(\lambda)$.

Матричные многочлены. На понятие λ-матрицы можно посмотреть с совершенно иной стороны. Назовем *матричным λ-многочленом порядка n над полем P* многочлен от λ, коэффициентами которого служат квадратные матрицы одного и того же порядка n с элементами из поля P; его общим видом будет

$$A_0\lambda^k+A_1\lambda^{k-1}+\dots+A_{k-1}\lambda+A_k. \tag{6}$$

Понимая, в соответствии с § 15, умножение матрицы A_i на λ^{k-i}, $i=0, 1, \dots, k$, как умножение на λ^{k-i} всех элементов матрицы A_i, а затем выполняя сложение матриц в соответствии с тем же § 15, мы получим, что *всякий матричный λ-многочлен порядка n можно записать в виде λ-матрицы порядка n.* Так,

$$\begin{pmatrix}4&0\\-1&1\end{pmatrix}\lambda^3+\begin{pmatrix}0&-3\\0&1\end{pmatrix}\lambda^2+\begin{pmatrix}1&2\\0&-2\end{pmatrix}\lambda+\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}4\lambda^3+\lambda&-3\lambda^2+2\lambda+1\\-\lambda^3&\lambda^3+\lambda^2-2\lambda\end{pmatrix}.$$

Обратно, *всякая λ-матрица порядка n может быть записана в виде матричного λ-многочлена порядка n.* Так,

$$\begin{pmatrix}3\lambda^2-5&\lambda+1\\\lambda^4+2\lambda&-3\end{pmatrix}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\lambda^4+\begin{pmatrix}3&0\\0&0\end{pmatrix}\lambda^2+\begin{pmatrix}0&1\\2&0\end{pmatrix}\lambda+\begin{pmatrix}-5&1\\0&-3\end{pmatrix}.$$

Соответствие между λ-матрицами и матричными λ-многочленами является взаимно однозначным и изоморфным в смысле § 46. Действительно, равенство λ-многочленов вида (6) как матриц равносильно равенству матричных коэффициентов при одинаковых степенях λ, а умножение матрицы на λ равносильно умножению ее на скалярную матрицу с λ на главной диагонали.

Пусть дана λ-матрица $A(\lambda)$, причем

$$A(\lambda) = A_0\lambda^k + A_1\lambda^{k-1} + \ldots + A_{k-1}\lambda + A_k,$$

где матрица A_0 не является нулевой. Число k назовем *степенью* λ-матрицы $A(\lambda)$; это будет, очевидно, наивысшая степень (по λ) элементов матрицы $A(\lambda)$.

Взгляд на λ-матрицы как на матричные многочлены позволяет развивать для λ-матриц теорию делимости, аналогичную теории делимости для числовых многочленов, но усложняемую, понятно, некоммутативностью умножения матриц и наличием делителей нуля. Мы ограничимся лишь вопросом об алгоритме деления с остатком.

Пусть над полем P даны λ-матрицы порядка n

$$A(\lambda) = A_0\lambda^k + A_1\lambda^{k-1} + \ldots + A_{k-1}\lambda + A_k,$$
$$B(\lambda) = B_0\lambda^l + B_1\lambda^{l-1} + \ldots + B_{l-1}\lambda + B_l,$$

причем предположим, что матрица B_0 невырожденная, т. е. существует матрица B_0^{-1}. Тогда над полем P можно найти такие λ-матрицы $Q_1(\lambda)$ и $R_1(\lambda)$ того же порядка n, что

$$A(\lambda) = B(\lambda)\,Q_1(\lambda) + R_1(\lambda), \tag{7}$$

причем степень $R_1(\lambda)$ меньше степени $B(\lambda)$ или же $R_1(\lambda) = 0$. С другой стороны, над полем P можно найти такие λ-матрицы $Q_2(\lambda)$ и $R_2(\lambda)$ порядка n, что

$$A(\lambda) = Q_2(\lambda)\,B(\lambda) + R_2(\lambda), \tag{8}$$

причем степень $R_2(\lambda)$ меньше степени $B(\lambda)$ или же $R_2(\lambda) = 0$. Матрицы $Q_1(\lambda)$ и $R_1(\lambda)$, а также $Q_2(\lambda)$ и $R_2(\lambda)$, удовлетворяющие этим условиям, определяются однозначно.

Доказательство этой теоремы проходит так же, как доказательство соответствующей теоремы для числовых многочленов (см. § 20). Пусть, например, условию (7) удовлетворяют также матрицы $\overline{Q}_1(\lambda)$ и $\overline{R}_1(\lambda)$, причем степень $\overline{R}_1(\lambda)$ меньше степени $B(\lambda)$. Тогда

$$B(\lambda)\left[Q_1(\lambda) - \overline{Q}_1(\lambda)\right] = \overline{R}_1(\lambda) - R_1(\lambda).$$

Степень правой части меньше l, степень же левой части, если квадратная скобка отлична от нуля, больше или равна l, так как матрица B_0 невырожденная. Отсюда следует единственность матриц $Q_1(\lambda)$ и $R_1(\lambda)$.

Для доказательства существования этих матриц заметим, что при $k \geqslant l$ степень разности

$$A(\lambda) - B(\lambda) \cdot B_0^{-1} A_0 \lambda^{k-l}$$

будет строго меньше k; поэтому $B_0^{-1} A_0 \lambda^{k-l}$ будет старшим членом матричного λ-многочлена $Q_1(\lambda)$. Дальше продолжается так же, как в § 20. С другой стороны, степень разности

$$A(\lambda) - A_0 B_0^{-1} \lambda^{k-l} \cdot B(\lambda)$$

также строго меньше k, т. е. $A_0 B_0^{-1} \lambda^{k-l}$ будет старшим членом матричного λ-многочлена $Q_2(\lambda)$. Мы видим, что λ-матрицы $Q_1(\lambda)$ и $Q_2(\lambda)$ (а также $R_1(\lambda)$ и $R_2(\lambda)$), удовлетворяющие условиям теоремы, действительно в общем случае будут различными.

Основная теорема о подобии матриц. Как уже отмечалось, у нас нет пока способа для решения вопроса, подобны ли данные числовые матрицы A и B (т. е. матрицы с элементами из основного поля P). С другой стороны, их характеристические матрицы $A - \lambda E$ и $B - \lambda E$ являются λ-матрицами и вопрос об эквивалентности этих матриц решается вполне эффективно. Легко понять поэтому, сколь велико значение следующей теоремы:

Матрицы A и B с элементами из поля P тогда и только тогда подобны, если их характеристические матрицы $A - \lambda E$ и $B - \lambda E$ эквивалентны.

В самом деле, пусть матрицы A и B подобны, т. е. над полем P существует такая невырожденная матрица C, что

$$B = C^{-1} A C.$$

Тогда

$$C^{-1}(A - \lambda E) C = C^{-1} A C - \lambda (C^{-1} E C) = B - \lambda E.$$

Невырожденные числовые матрицы C^{-1} и C являются, однако, унимодулярными λ-матрицами. Мы видим, что матрица $B - \lambda E$ получена умножением матрицы $A - \lambda E$ слева и справа на унимодулярные матрицы, т. е. $A - \lambda E \sim B - \lambda E$.

Доказательство обратного утверждения является более сложным. Пусть

$$A - \lambda E \sim B - \lambda E.$$

Тогда существуют такие унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$, что

$$U(\lambda)(A - \lambda E) V(\lambda) = B - \lambda E. \tag{9}$$

Учитывая, что для унимодулярных матриц обратные матрицы существуют и являются λ-матрицами, выведем из (9) следующие равенства, используемые ниже:

$$\left.\begin{aligned} U(\lambda)(A - \lambda E) &= (B - \lambda E) V^{-1}(\lambda), \\ (A - \lambda E) V(\lambda) &= U^{-1}(\lambda)(B - \lambda E). \end{aligned}\right\} \tag{10}$$

Так как λ-матрица $B-\lambda E$ имеет по λ степень 1, причем старшим коэффициентом соответствующего матричного многочлена служит невырожденная матрица $-E$, то к матрицам $U(\lambda)$ и $B-\lambda E$ можно применить алгоритм деления с остатком: существуют такие матрицы $Q_1(\lambda)$ и R_1 — последняя, если она отлична от нуля, должна иметь по λ степень 0, т. е. от λ не зависит, — что

$$U(\lambda)=(B-\lambda E)\,Q_1(\lambda)+R_1. \tag{11}$$

Аналогично

$$V(\lambda)=Q_2(\lambda)(B-\lambda E)+R_2. \tag{12}$$

Используя (11) и (12), из (9) получаем:

$$R_1(A-\lambda E)R_2=(B-\lambda E)-U(\lambda)(A-\lambda E)\,Q_2(\lambda)(B-\lambda E)-$$
$$-(B-\lambda E)\,Q_1(\lambda)(A-\lambda E)\,V(\lambda)+(B-\lambda E)\,Q_1(\lambda)(A-\lambda E)\,Q_2(\lambda)(B-\lambda E)$$

или, ввиду (10),

$$R_1(A-\lambda E)R_2=(B-\lambda E)-(B-\lambda E)\,V^{-1}(\lambda)\,Q_2(\lambda)(B-\lambda E)-$$
$$-(B-\lambda E)\,Q_1(\lambda)\,U^{-1}(\lambda)(B-\lambda E)+$$
$$+(B-\lambda E)\,Q_1(\lambda)(A-\lambda E)\,Q_2(\lambda)(B-\lambda E)=(B-\lambda E)\times$$
$$\times\{E-[V^{-1}(\lambda)\,Q_2(\lambda)+Q_1(\lambda)\,U^{-1}(\lambda)-Q_1(\lambda)(A-\lambda E)\,Q_2(\lambda)](B-\lambda E)\}.$$

Квадратная скобка, стоящая справа, равна в действительности нулю: в противном случае она, являясь λ-матрицей, так как и $V^{-1}(\lambda)$, и $U^{-1}(\lambda)$ суть λ-матрицы, имела бы по меньшей мере степень 0, а тогда степень фигурной скобки была бы не меньше 1 и, следовательно, степень всей правой части была бы не меньше 2. Это, однако, невозможно, так как слева стоит λ-матрица степени 1.

Таким образом,

$$R_1(A-\lambda E)R_2=B-\lambda E,$$

откуда, приравнивая матричные коэффициенты при одинаковых степенях λ, получаем

$$R_1AR_2=B, \tag{13}$$
$$R_1R_2=E. \tag{14}$$

Равенство (14) показывает, что числовая матрица R_2 не только отлична от нуля, но даже является невырожденной, причем

$$R_2^{-1}=R_1,$$

а тогда равенство (13) принимает вид

$$R_2^{-1}AR_2=B,$$

что и доказывает подобие матриц A и B.

Одновременно мы научились находить ту невырожденную матрицу R_2, которая трансформирует матрицу A в матрицу B. Именно, если матрицы $A-\lambda E$ и $B-\lambda E$ эквивалентны, то первая конечным числом элементарных преобразований переводится во вторую. Берем те из этих преобразований, которые относятся к столбцам, и произведение соответствующих элементарных матриц, взятых в том же порядке, обозначаем через $V(\lambda)$. Делим затем $V(\lambda)$ на $B-\lambda E$, причем так, чтобы частное стояло слева от делителя (см. (8)). Остаток от этого деления и будет матрицей R_2.

Указанное деление можно на самом деле не выполнять, а воспользоваться следующей леммой, которая найдет применение также в § 62:

Лемма. *Пусть*

$$V(\lambda)=V_0\lambda^s+V_1\lambda^{s-1}+\ldots+V_{s-1}\lambda+V_s,\quad V_0\neq 0. \tag{15}$$

Если

$$V(\lambda)=(\lambda E-B)\,Q_1(\lambda)+R_1, \tag{16}$$
$$V(\lambda)=Q_2(\lambda)\,(\lambda E-B)+R_2,$$

то

$$R_1=B^sV_0+B^{s-1}V_1+\ldots+BV_{s-1}+V_s, \tag{17}$$
$$R_2=V_0B^s+V_1B^{s-1}+\ldots+V_{s-1}B+V_s.$$

Достаточно доказать хотя бы первое из двух утверждений леммы — второе доказывается вполне аналогично. Доказательство состоит в непосредственной проверке справедливости равенства (16), если многочлен $V(\lambda)$ будет заменен его записью (15), вместо R_1 будет подставлено (17), а в качестве $Q_1(\lambda)$ будет взят многочлен

$$Q_1(\lambda)=V_0\lambda^{s-1}+(BV_0+V_1)\lambda^{s-2}+(B^2V_0+BV_1+V_2)\lambda^{s-3}+\ldots$$
$$\ldots+(B^{s-1}V_0+B^{s-2}V_1+\ldots+V_{s-1}).$$

Проверка эта предоставляется читателю.

Пример. Даны матрицы

$$A=\begin{pmatrix}-2 & 1\\ 0 & 3\end{pmatrix},\quad B=\begin{pmatrix}-10 & -4\\ 26 & 11\end{pmatrix}.$$

Их характеристические матрицы эквивалентны, так как приводятся к одному и тому же каноническому виду

$$\begin{pmatrix}1 & 0\\ 0 & \lambda^2-\lambda-6\end{pmatrix},$$

поэтому матрицы A и B подобны.

Для разыскания матрицы R_2, трансформирующей A в B, найдем какую-либо цепочку элементарных преобразований, переводящую $A-\lambda E$ в $B-\lambda E$. Так,

$$A-\lambda E=\begin{pmatrix}-2-\lambda & 1\\ 0 & 3-\lambda\end{pmatrix}\sim\begin{pmatrix}-2-\lambda & 1\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}8+4\lambda & -4\\ -16-8\lambda & 11-\lambda\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}40+4\lambda & -4\\ -104 & 11-\lambda\end{pmatrix}\sim\begin{pmatrix}-10-\lambda & -4\\ 26 & 11-\lambda\end{pmatrix}=B-\lambda E.$$

К столбцам относятся два последних преобразования: к первому столбцу прибавляется второй, умноженный на -8, а затем первый столбец умножается на $-\frac{1}{4}$. Произведение соответствующих элементарных матриц будет

$$V(\lambda)=\begin{pmatrix}1 & 0\\ -8 & 1\end{pmatrix}\begin{pmatrix}-\frac{1}{4} & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}-\frac{1}{4} & 0\\ 2 & 1\end{pmatrix}.$$

Эта матрица не зависит от λ и поэтому она и будет искомой матрицей R_2.

Конечно, матрица, трансформирующая A в B, определяется далеко не однозначно. Такой будет, например, также матрица

$$\begin{pmatrix}3 & 1\\ 2 & 1\end{pmatrix}.$$

§ 61. Жорданова нормальная форма

Будем рассматривать сейчас квадратные матрицы порядка n с элементами из поля P. Будет выделен один специальный тип таких матриц, так называемые ж о р д а н о в ы матрицы, и будет показано, что эти матрицы служат нормальной формой для весьма широкого класса матриц. Именно, *матрицы, все характеристические корни которых лежат в основном поле P (и только такие матрицы), подобны некоторым жордановым матрицам, т. е., как говорят, они приводятся к жордановой нормальной форме.* Отсюда будет следовать, если в качестве поля P взято поле комплексных чисел, что *всякая матрица с комплексными элементами приводится в поле комплексных чисел к жордановой нормальной форме.*

Введем необходимые определения. *Жордановой клеткой* порядка k, относящейся к числу λ_0, называется матрица порядка k, $1\leqslant k\leqslant n$, имеющая вид

$$\begin{pmatrix}\lambda_0 & 1 & & & \mathbf{0}\\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ \mathbf{0} & & & & \lambda_0\end{pmatrix}; \qquad (1)$$

иными словами, на ее главной диагонали стоит одно и то же число λ_0 из поля P; параллель, ближайшая к главной диагонали сверху, сплошь занята числом 1; все остальные элементы матрицы равны нулю. Так,

$$(\lambda_0), \quad \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}$$

будут соответственно жордановыми клетками первого, второго и третьего порядков.

Жордановой матрицей порядка n называется матрица порядка n, имеющая вид

$$J = \begin{pmatrix} \boxed{J_1} & & & & 0 \\ & \boxed{J_2} & & & \\ & & \ddots & & \\ 0 & & & & \boxed{J_s} \end{pmatrix}; \tag{2}$$

здесь вдоль главной диагонали идут жордановы клетки $J_1, J_2, \ldots, J_s$ некоторых порядков, не обязательно различных, относящиеся к некоторым числам из поля P, также не обязательно различным; все места вне этих клеток заняты нулями. При этом $s \geqslant 1$, т. е. одна жорданова клетка порядка n принадлежит к числу жордановых матриц этого порядка, и, понятно, $s \leqslant n$.

Заметим, хотя это и не будет дальше использоваться, что строение жордановой матрицы можно было бы описать, не прибегая к понятию жордановой клетки. Очевидно, именно, что матрица J будет жордановой матрицей тогда и только тогда, если она имеет вид

$$\begin{pmatrix} \lambda_1 & \varepsilon_1 & & & 0 \\ & \lambda_2 & \varepsilon_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \varepsilon_{n-1} \\ 0 & & & & \lambda_n \end{pmatrix},$$

где λ_i, $i = 1, 2, \ldots, n$,— произвольные числа из поля P, а каждое ε_j, $j = 1, 2, \ldots, n-1$, равно единице или нулю, причем, если $\varepsilon_j = 1$, то $\lambda_j = \lambda_{j+1}$.

Диагональные матрицы являются частным случаем жордановых матриц: это будут в точности те жордановы матрицы, у которых все жордановы клетки имеют порядок 1.

Нашей ближайшей целью является разыскание канонического вида для характеристической матрицы $J - \lambda E$

произвольной жордановой матрицы J порядка n. Найдем сначала канонический вид для характеристической матрицы

$$\begin{pmatrix} \lambda_0-\lambda & 1 & & & 0 \\ & \lambda_0-\lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_0-\lambda \end{pmatrix} \quad (3)$$

одной жордановой клетки (1) порядка k. Вычисляя определитель этой матрицы и вспоминая, что старший коэффициент многочлена $d_k(\lambda)$ должен равняться 1, получаем, что

$$d_k(\lambda) = (\lambda-\lambda_0)^k.$$

С другой стороны, среди миноров $(k-1)$-го порядка матрицы (3) имеется минор, равный единице, а именно тот, который получается после вычеркивания первого столбца и последней строки этой матрицы. Поэтому

$$d_{k-1}(\lambda) = 1.$$

Отсюда следует, что *каноническим видом для матрицы* (3) *служит следующая λ-матрица порядка k*:

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & (\lambda-\lambda_0)^k \end{pmatrix}. \quad (4)$$

Докажем теперь следующую лемму:

Если многочлены $\varphi_1(\lambda), \varphi_2(\lambda), \ldots, \varphi_t(\lambda)$ из кольца $P[\lambda]$ попарно взаимно просты, то имеет место следующая эквивалентность:

$$\begin{pmatrix} \varphi_1(\lambda) & & & 0 \\ & \varphi_2(\lambda) & & \\ & & \ddots & \\ 0 & & & \varphi_t(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \prod\limits_{i=1}^{t} \varphi_i(\lambda) \end{pmatrix}.$$

Достаточно, очевидно, рассмотреть случай $t=2$. Так как многочлены $\varphi_1(\lambda)$ и $\varphi_2(\lambda)$ взаимно просты, то в кольце $P[\lambda]$ существуют такие многочлены $u_1(\lambda)$ и $u_2(\lambda)$, что

$$\varphi_1(\lambda)\, u_1(\lambda) + \varphi_2(\lambda)\, u_2(\lambda) = 1.$$

Поэтому

$$\begin{pmatrix}\varphi_1(\lambda) & 0\\ 0 & \varphi_2(\lambda)\end{pmatrix}\sim\begin{pmatrix}\varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda)\\ 0 & \varphi_2(\lambda)\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}\varphi_1(\lambda) & \varphi_1(\lambda)u_1(\lambda)+\varphi_2(\lambda)u_2(\lambda)\\ 0 & \varphi_2(\lambda)\end{pmatrix}=\begin{pmatrix}\varphi_1(\lambda) & 1\\ 0 & \varphi_2(\lambda)\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}1 & \varphi_1(\lambda)\\ \varphi_2(\lambda) & 0\end{pmatrix}\sim\begin{pmatrix}1 & \varphi_1(\lambda)\\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda)\end{pmatrix}\sim$$

$$\sim\begin{pmatrix}1 & 0\\ 0 & -\varphi_1(\lambda)\varphi_2(\lambda)\end{pmatrix}\sim\begin{pmatrix}1 & 0\\ 0 & \varphi_1(\lambda)\varphi_2(\lambda)\end{pmatrix},$$

что и требовалось доказать.

Перейдем теперь к рассмотрению характеристической матрицы

$$J-\lambda E=\begin{pmatrix}\boxed{J_1-\lambda E_1} & & & \mathbf{0}\\ & \boxed{J_2-\lambda E_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{J_s-\lambda E_s}\end{pmatrix} \tag{5}$$

для жордановой матрицы J вида (2); здесь E_i, $i=1, 2, \ldots, s$, есть единичная матрица того же порядка, что и клетка J_i. Пусть жордановы клетки матрицы J относятся к следующим различным числам: $\lambda_1, \lambda_2, \ldots, \lambda_t$, где $t\leqslant s$. Пусть, далее, к числу λ_i, $i=1, 2, \ldots, t$, относится q_i жордановых клеток, $q_i\geqslant 1$, и пусть порядки этих клеток, расположенные в невозрастающем порядке, будут

$$k_{i1}\geqslant k_{i2}\geqslant\ldots\geqslant k_{iq_i}. \tag{6}$$

Отметим, хотя и не будем этим пользоваться, что

$$\sum_{i=1}^{t} q_i=s,$$

$$\sum_{i=1}^{t}\sum_{j=1}^{q_i} k_{ij}=n.$$

Применяя элементарные преобразования к тем строкам и столбцам матрицы (5), которые проходят через клетку $J_i-\lambda E_i$ этой матрицы, мы не будем затрагивать, очевидно, других диагональных клеток. Отсюда следует, что в матрице (5) можно при помощи элементарных преобразований заменить каждую клетку $J_i-\lambda E_i$, $i=1, 2, \ldots, s$, соответствующей клеткой вида (4). Иными словами, *матрица $J-\lambda E$ эквивалентна диагональной матрице, на диагонали которой стоят, помимо некоторого числа единиц, также*

следующие многочлены, соответствующие всем жордановым клеткам матрицы J:

$$\left.\begin{array}{llll}(\lambda-\lambda_1)^{k_{11}}, & (\lambda-\lambda_1)^{k_{12}}, & \ldots, & (\lambda-\lambda_1)^{k_{1q_1}},\\ (\lambda-\lambda_2)^{k_{21}}, & (\lambda-\lambda_2)^{k_{22}}, & \ldots, & (\lambda-\lambda_2)^{k_{2q_2}},\\ \cdots & \cdots & \cdots & \cdots\\ (\lambda-\lambda_t)^{k_{t1}}, & (\lambda-\lambda_t)^{k_{t2}}, & \ldots, & (\lambda-\lambda_t)^{k_{tq_t}}.\end{array}\right\} \qquad (7)$$

Мы не указываем при этом те места на диагонали, на которых стоят многочлены (7), так как в любой диагональной λ-матрице диагональные элементы можно произвольно переставлять при помощи перестановок строк и одноименных столбцов. Это замечание следует учитывать и в дальнейшем.

Пусть q — наибольшее среди чисел q_i, $i=1, 2, \ldots, t$. Обозначим через $e_{n-j+1}(\lambda)$ произведение многочленов, стоящих в j-м столбце таблицы (7), $j=1, 2, \ldots, q$, т. е.

$$e_{n-j+1}(\lambda)=\prod_{i=1}^{t}(\lambda-\lambda_i)^{k_{ij}}; \qquad (8)$$

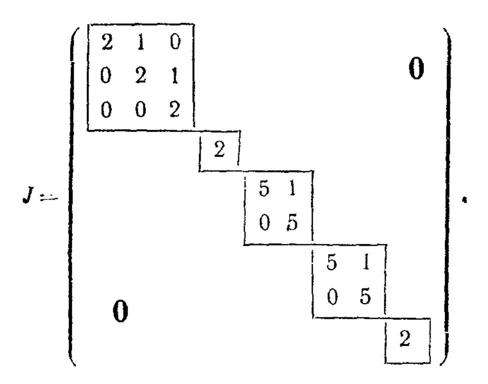
если при этом в j-м столбце имеются пустые места — для некоторых i может оказаться, что $q_i<j$, — то соответствующие множители в (8) считаем равными единице. Так как числа $\lambda_1, \lambda_2, \ldots, \lambda_t$ по условию различные, то степени линейных двучленов, стоящие в j-м столбце таблицы (7), попарно взаимно просты. Поэтому, на основании доказанной выше леммы, они при помощи элементарных преобразований могут быть заменены в рассматриваемой диагональной матрице их произведением $e_{n-j+1}(\lambda)$ и некоторым числом единиц.

Проделав это для $j=1, 2, \ldots, q$, мы получим, что

$$J-\lambda E \sim \begin{pmatrix} 1 & & & & & & & 0\\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & e_{n-q+1}(\lambda) & & & & \\ & & & & \ddots & & & \\ & & & & & & e_{n-1}(\lambda) & \\ 0 & & & & & & & e_n(\lambda)\end{pmatrix}. \qquad (9)$$

Это и будет искомый канонический вид матрицы $J-\lambda E$. Действительно, старшие коэффициенты всех многочленов, стоящих в (9) на главной диагонали, равны единице и каждый из этих многочленов нацело делится на предыдущий ввиду условия (6).

Пример. Пусть



Для этой жордановой матрицы 9-го порядка таблица многочленов (7) имеет вид

$$(\lambda-2)^3, \quad \lambda-2, \quad \lambda-2,$$
$$(\lambda-5)^2, \quad (\lambda-5)^2.$$

Поэтому инвариантными множителями матрицы J будут многочлены

$$e_9(\lambda)=(\lambda-2)^3(\lambda-5)^2,$$
$$e_8(\lambda)=(\lambda-2)(\lambda-5)^2,$$
$$e_7(\lambda)=\lambda-2,$$

в то время как $e_6(\lambda)=\ldots=e_1(\lambda)=1$.

Теперь, когда мы научились по виду данной жордановой матрицы J сразу писать канонический вид ее характеристической матрицы, может быть доказана следующая теорема:

Две жордановы матрицы тогда и только тогда подобны, если они состоят из одних и тех же жордановых клеток, т. е. отличаются, быть может, лишь расположением этих клеток вдоль главной диагонали.

В самом деле, таблица многочленов (7) полностью определялась набором жордановых клеток жордановой матрицы J и в ней никак не отражалось расположение жордановых клеток вдоль главной диагонали этой матрицы. Отсюда следует, что если жордановы матрицы J и J' обладают одним и тем же набором жордановых клеток, то им соответствует одна и та же таблица многочленов (7), а поэтому одни и те же многочлены (8). Таким образом, характеристические матрицы $J-\lambda E$ и $J'-\lambda E$ обладают одинаковыми инвариантными множителями, т. е. эквивалентны, а поэтому сами матрицы J и J' подобны.

Обратно, если жордановы матрицы J и J' подобны, то их характеристические матрицы обладают одинаковыми инвариантными множителями. Пусть многочлены (8) для $j=1, 2, \ldots, q$ будут те из этих инвариантных множителей, которые отличны от единицы. Однако по многочленам (8) восстанавливается таблица многочленов (7). Именно, многочлены (8) разлагаются в произведение степеней линейных множителей, так как этим свойством обладают, как уже доказано, инвариантные множители характеристической матрицы для любой жордановой матрицы. Таблица (7) как раз и состоит из всех тех максимальных степеней линейных множителей, на которые разлагаются многочлены (8). Наконец, по таблице (7) восстанавливаются жордановы клетки исходных жордановых матриц: каждому многочлену $(\lambda-\lambda_i)^{k_{ij}}$ из таблицы (7) соответствует жорданова клетка порядка k_{ij}, относящаяся к числу λ_i. Этим доказано, что матрицы J и J' состоят из одних и тех же жордановых клеток и отличаются, быть может, лишь их расположением.

Из этой теоремы следует, в частности, *что жорданова матрица, подобная диагональной матрице, сама диагональна и что две диагональные матрицы тогда и только тогда подобны, если получаются друг из друга перестановкой чисел, стоящих на главной диагонали.*

Приведение матрицы к жордановой нормальной форме. Если матрица A с элементами из поля P *приводится к жордановой нормальной форме,* т. е. подобна жордановой матрице, то, как следует из доказанной выше теоремы, *жорданова нормальная форма определяется для матрицы A однозначно с точностью до расположения жордановых клеток на главной диагонали.* Условие для того, чтобы матрица A допускала такое приведение, указывается в следующей теореме, доказательство которой дает одновременно практический способ для разыскания жордановой матрицы, подобной матрице A, если такая жорданова матрица существует. При этом заметим, что приводимость в поле P означает, что все элементы трансформирующей матрицы содержатся в поле P.

Матрица A с элементами из поля P тогда и только тогда приводится в поле P к жордановой нормальной форме, если все характеристические корни матрицы A лежат в самом основном поле P.

В самом деле, если матрица A подобна жордановой матрице J, то эти две матрицы обладают одними и теми же характеристическими корнями. Характеристические корни матрицы J находятся, однако, без всяких затруднений: так как определитель матрицы $J-\lambda E$ равен произведению ее элементов, стоящих на главной диагонали, то многочлен $|J-\lambda E|$ разлагается над полем P на линейные множители и его корнями служат числа, стоящие на главной диагонали матрицы J, и только они.

Обратно, пусть все характеристические корни матрицы A лежат в самом поле P. Если отличные от 1 инвариантные множители матрицы $A-\lambda E$ будут

$$e_{n-q+1}(\lambda),\ \ldots,\ e_{n-1}(\lambda),\ e_n(\lambda), \tag{10}$$

то

$$|A-\lambda E|=(-1)^n e_{n-q+1}(\lambda)\ \ldots\ e_{n-1}(\lambda)\, e_n(\lambda).$$

Действительно, определители матрицы $A-\lambda E$ и ее канонической матрицы могут отличаться друг от друга лишь постоянным множителем, который на самом деле равен $(-1)^n$, так как именно таков старший коэффициент характеристического многочлена $|A-\lambda E|$. Таким образом, среди многочленов (10) нет равных нулю, сумма степеней этих многочленов равна n и все они разлагаются над полем P на линейные множители — последнее ввиду того, что, по условию, многочлен $|A-\lambda E|$ обладает таким разложением.

Пусть (8) будут разложения многочленов (10) в произведения степеней линейных множителей. Назовем *элементарными делителями многочлена* e_{n-j+1}, $j=1, 2, \ldots, q$, отличные от единицы степени различных линейных двучленов, входящие в его разложение (8), т. е.

$$(\lambda-\lambda_1)^{k_{1j}},\ (\lambda-\lambda_2)^{k_{2j}},\ \ldots,\ (\lambda-\lambda_t)^{k_{tj}}.$$

Элементарные делители всех многочленов (10) назовем *элементарными делителями матрицы* A и выпишем их в виде таблицы (7).

Возьмем теперь жорданову матрицу J порядка n, составленную из жордановых клеток, определяемых следующим образом: к а ж д о м у э л е м е н т а р н о м у д е л и т е л ю $(\lambda-\lambda_i)^{k_{ij}}$ м а т р и ц ы A с т а в и м в с о о т в е т с т в и е ж о р д а н о в у к л е т к у п о р я д к а k_{ij}, о т н о с я щ у ю с я к ч и с л у λ_i. Очевидно, что отличными от 1 инвариантными множителями матрицы $J-\lambda E$ будут многочлены (10) и только они. Поэтому матрицы $A-\lambda E$ и $J-\lambda E$ эквивалентны и, следовательно, матрица A подобна жордановой матрице J.

П р и м е р. Пусть дана матрица

$$A=\begin{pmatrix} -16 & -17 & 87 & -108 \\ 8 & 9 & -42 & 54 \\ -3 & -3 & 16 & -18 \\ -1 & -1 & 6 & -8 \end{pmatrix}.$$

Приводя обычным способом матрицу $A-\lambda E$ к каноническому виду, получим, что отличными от единицы инвариантными множителями этой матрицы будут многочлены

$$e_4(\lambda)=(\lambda-1)^2(\lambda+2),$$
$$e_3(\lambda)=\lambda-1.$$

Мы видим, что матрица A приводится к жордановой нормальной форме даже в поле рациональных чисел. Ее элементарными делителями являются много-

члены $(\lambda-1)^2$, $\lambda-1$ и $\lambda+2$, а поэтому жордановой нормальной формой матрицы A служит матрица

$$J=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-2\end{pmatrix}.$$

Если бы мы хотели найти ту невырожденную матрицу, которая трансформирует матрицу A в матрицу J, то должны были бы воспользоваться указаниями, сделанными в конце предшествующего параграфа.

На основании предшествующих результатов может быть доказано, наконец, следующее **необходимое и достаточное условие приводимости матрицы к диагональному виду**, условие, из которого немедленно вытекает достаточный критерий приводимости к диагональному виду, доказанный в § 33.

Матрица A порядка n с элементами из поля P тогда и только тогда приводится к диагональному виду, если все корни последнего инвариантного множителя $e_n(\lambda)$ ее характеристической матрицы лежат в поле P, причем среди этих корней нет кратных.

В самом деле, приводимость матрицы к диагональному виду равносильна приводимости к такому жорданову виду, все жордановы клетки которого имеют порядок 1. Иными словами, все элементарные делители матрицы A должны быть многочленами первой степени. Так как, однако, все инвариантные множители матрицы $A-\lambda E$ являются делителями многочлена $e_n(\lambda)$, то последнее условие равносильно тому, что все элементарные делители многочлена $e_n(\lambda)$ имеют степень 1, что и требовалось доказать.

§ 62. Минимальный многочлен

Пусть дана квадратная матрица A порядка n с элементами из поля P. Если

$$f(\lambda)=\alpha_0\lambda^k+\alpha_1\lambda^{k-1}+\ldots+\alpha_{k-1}\lambda+\alpha_k$$

— произвольный многочлен из кольца $P[\lambda]$, то матрица

$$f(A)=\alpha_0A^k+\alpha_1A^{k-1}+\ldots+\alpha_{k-1}A+\alpha_kE$$

будет называться *значением* многочлена $f(\lambda)$ при $\lambda=A$; обращаем внимание на то, что свободный член многочлена $f(\lambda)$ умножается при этом на нулевую степень матрицы A, т. е. на единичную матрицу E.

Легко проверяется, что если

$$f(\lambda)=\varphi(\lambda)+\psi(\lambda)$$

или

$$f(\lambda)=u(\lambda)\,v(\lambda),$$

то

$$f(A)=\varphi(A)+\psi(A)$$

и, соответственно,

$$f(A)=u(A)\,v(A).$$

Если многочлен $f(\lambda)$ аннулируется матрицей A, т. е.

$$f(A)=0,$$

то матрицу A будем называть *матричным корнем* или, там, где это не может вызвать недоразумений, просто *корнем* многочлена $f(\lambda)$.

Всякая матрица A служит корнем некоторого ненулевого многочлена.

Мы знаем, в самом деле, что все квадратные матрицы порядка n составляют над полем P n^2-мерное векторное пространство. Отсюда следует, что система n^2+1 матриц

$$A^{n^2},\ A^{n^2-1},\ \ldots,\ A,\ E$$

линейно зависима над полем P, т. е. в P существуют такие элементы $\alpha_0,\ \alpha_1,\ \ldots,\ \alpha_{n^2},\ \alpha_{n^2+1}$, не все равные нулю, что

$$\alpha_0 A^{n^2}+\alpha_1 A^{n^2-1}+\ldots+\alpha_{n^2}A+\alpha_{n^2+1}E=0.$$

Таким образом, матрица A оказалась корнем ненулевого многочлена

$$\varphi(\lambda)=\alpha_0\lambda^{n^2}+\alpha_1\lambda^{n^2-1}+\ldots+\alpha_{n^2}\lambda+\alpha_{n^2+1},$$

степень которого не превосходит n^2.

Матрица A служит корнем и для некоторых таких многочленов, старшие коэффициенты которых равны единице — достаточно взять любой ненулевой многочлен, аннулируемый матрицей A, и разделить этот многочлен на его старший коэффициент. Многочлен наименьшей степени со старшим коэффициентом 1, аннулируемый матрицей A, называется *минимальным многочленом* матрицы A. Заметим, что *минимальный многочлен матрицы A определен однозначно*, так как разность двух таких многочленов имела бы меньшую степень, чем каждый из них, но также аннулировалась бы матрицей A.

Всякий многочлен $f(\lambda)$, аннулируемый матрицей A, делится нацело на минимальный многочлен $m(\lambda)$ этой матрицы.

В самом деле, если

$$f(\lambda)=m(\lambda)\,q(\lambda)+r(\lambda),$$

где степень $r(\lambda)$ меньше степени $m(\lambda)$, то

$$f(A)=m(A)\,q(A)+r(A)$$

и из $f(A)=m(A)=0$ следует $r(A)=0$, что противоречит определению минимального многочлена.

Докажем теперь следующую теорему:

Минимальный многочлен матрицы A совпадает с последним инвариантным множителем $e_n(\lambda)$ характеристической матрицы $A-\lambda E$.

Доказательство. Сохраняя обозначения и используя результаты § 59, можно написать равенство

$$(-1)^n |A-\lambda E| = d_{n-1}(\lambda)\, e_n(\lambda). \tag{1}$$

Отсюда следует, в частности, что многочлены $e_n(\lambda)$ и $d_{n-1}(\lambda)$ не будут нулевыми. Обозначим, далее, через $B(\lambda)$ присоединенную матрицу к матрице $A-\lambda E$ (см. § 14),

$$B(\lambda) = (A-\lambda E)^*.$$

Как вытекает из § 14, равенство (3), справедливо равенство

$$(A-\lambda E)\, B(\lambda) = |A-\lambda E|\, E. \tag{2}$$

С другой стороны, так как элементами матрицы $B(\lambda)$ служат взятые со знаками плюс или минус миноры $(n-1)$-го порядка матрицы $A-\lambda E$ и только они, а многочлен $d_{n-1}(\lambda)$ есть общий наибольший делитель всех этих миноров, то

$$B(\lambda) = d_{n-1}(\lambda)\, C(\lambda), \tag{3}$$

причем **наибольший общий делитель элементов матрицы $C(\lambda)$ равен 1.**

Из равенств (2), (3) и (1) вытекает равенство

$$(A-\lambda E)\, d_{n-1}(\lambda)\, C(\lambda) = (-1)^n d_{n-1}(\lambda)\, e_n(\lambda)\, E.$$

Это равенство можно сократить на ненулевой множитель $d_{n-1}(\lambda)$, как вытекает из следующего общего замечания: если $\varphi(\lambda)$ — ненулевой многочлен, $D(\lambda) = (d_{ij}(\lambda))$ — ненулевая λ-матрица, причем пусть $d_{st}(\lambda) \neq 0$, то в матрице $\varphi(\lambda)\, D(\lambda)$ на месте (s, t) будет стоять отличный от нуля элемент $\varphi(\lambda)\, d_{st}(\lambda)$. Таким образом,

$$(A-\lambda E)\, C(\lambda) = (-1)^n e_n(\lambda)\, E,$$

откуда

$$e_n(\lambda)\, E = (\lambda E - A)\, [(-1)^{n+1} C(\lambda)]. \tag{4}$$

Это равенство показывает, что остаток от «левого» деления λ-матрицы, стоящей слева, на двучлен $\lambda E - A$ равен нулю. Из леммы, доказанной в конце § 60, вытекает, однако, что этот остаток равен матрице $e_n(A)\, E = e_n(A)$. Действительно, матрица $e_n(\lambda)\, E$ может быть записана как матричный λ-многочлен, коэффициенты которого являются скалярными матрицами, т. е. перестановочны с матрицей A. Таким образом

$$e_n(A) = 0,$$

т. е. многочлен $e_n(\lambda)$ действительно аннулируется матрицей A.

Отсюда следует, что многочлен $e_n(\lambda)$ нацело делится на минимальный многочлен $m(\lambda)$ матрицы A,

$$e_n(\lambda) = m(\lambda)\, q(\lambda). \qquad (5)$$

Ясно, что старший коэффициент многочлена $q(\lambda)$ равен единице.

Так как $m(A) = 0$, то снова, ввиду той же леммы из § 60, остаток от левого деления λ-матрицы $m(\lambda)E$ на двучлен $\lambda E - A$ равен нулю, т. е.

$$m(\lambda)E = (\lambda E - A)\, Q(\lambda). \qquad (6)$$

Равенства (5), (4) и (6) приводят к равенству

$$(\lambda E - A)\,[(-1)^{n+1} C(\lambda)] = (\lambda E - A)\,[Q(\lambda)\, q(\lambda)].$$

Обе части этого равенства можно сократить на общий множитель $\lambda E - A$, так как старший коэффициент E этого матричного λ-многочлена является невырожденной матрицей. Таким образом,

$$C(\lambda) = (-1)^{n+1} Q(\lambda)\, q(\lambda).$$

Мы помним, однако, что наибольший общий делитель элементов матрицы $C(\lambda)$ равен 1. Поэтому многочлен $q(\lambda)$ должен иметь нулевую степень, а так как его старший коэффициент равен 1, то $q(\lambda) = 1$. Таким образом, ввиду (5),

$$e_n(\lambda) = m(\lambda),$$

что и требовалось доказать.

Так как, ввиду (1), характеристический многочлен матрицы A нацело делится на многочлен $e_n(\lambda)$, то из доказанной сейчас теоремы вытекает следующая

Теорема Гамильтона—Кэли. *Всякая матрица является корнем своего характеристического многочлена.*

Минимальный многочлен линейного преобразования. Докажем сначала следующее утверждение:

Если матрицы A и B подобны и если многочлен $f(\lambda)$ аннулируется матрицей A, то он аннулируется и матрицей B.

Действительно, пусть

$$B = C^{-1}AC.$$

Если

$$f(\lambda) = \alpha_0\lambda^k + \alpha_1\lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k,$$

то

$$\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1}A + \alpha_k E = 0.$$

Трансформируя обе части этого равенства матрицей C, получаем:

$$C^{-1}(\alpha_0 A^k + \alpha_1 A^{k-1} + \ldots + \alpha_{k-1}A + \alpha_k E)\,C =$$
$$= \alpha_0 (C^{-1}AC)^k + \alpha_1 (C^{-1}AC)^{k-1} + \ldots + \alpha_{k-1}(C^{-1}AC) + \alpha_k E =$$
$$= \alpha_0 B^k + \alpha_1 B^{k-1} + \ldots + \alpha_{k-1}B + \alpha_k E = 0,$$

т. е. $f(B) = 0$.

Отсюда следует, что *подобные матрицы обладают одним и тем же минимальным многочленом.*

Пусть теперь φ будет линейное преобразование n-мерного линейного пространства над полем P. Матрицы, задающие это преобразование в разных базах пространства, подобны между собой. Общий минимальный многочлен этих матриц называется *минимальным многочленом линейного преобразования* φ.

Используя операции над линейными преобразованиями, введенные в § 32, можно ввести понятие *значения* многочлена

$$f(\lambda) = \alpha_0\lambda^k + \alpha_1\lambda^{k-1} + \ldots + \alpha_{k-1}\lambda + \alpha_k$$

из кольца $P[\lambda]$ при λ, равном линейному преобразованию φ: это будет линейное преобразование

$$f(\varphi) = \alpha_0\varphi^k + \alpha_1\varphi^{k-1} + \ldots + \alpha_{k-1}\varphi + \alpha_k\varepsilon,$$

где ε — тождественное преобразование.

Мы скажем, далее, что многочлен $f(\lambda)$ *аннулируется* линейным преобразованием φ, если

$$f(\varphi) = \omega,$$

где ω — нулевое преобразование.

Учитывая связь между операциями над линейными преобразованиями и над матрицами, читатель без труда докажет, что *минимальный многочлен линейного преобразования* φ *является тем однозначно определенным многочленом наименьшей степени со старшим коэффициентом* 1, *который аннулируется преобразованием* φ. После этого результаты, полученные выше, в частности теорема Гамильтона—Кэли, могут быть переформулированы на языке линейных преобразований.

ГЛАВА ЧЕТЫРНАДЦАТАЯ

ГРУППЫ

§ 63. Определение и примеры групп

Кольца и поля, игравшие столь большую роль в предшествующих главах, являются алгебраическими системами с двумя независимыми операциями: сложением и умножением. В различных отделах математики и в ее приложениях весьма часто встречаются, однако, и такие алгебраические системы, в которых определена лишь одна алгебраическая операция. Так, ограничиваясь пока примерами, уже появлявшимися в нашей книге, отметим, что в множестве подстановок n-й степени (см. § 3) нами была определена лишь одна операция — умножение. С другой стороны, в определение векторного пространства (§ 8) входит сложение векторов, в то время как умножение векторов не было нами определено (заметим, что умножение вектора на число не удовлетворяет данному в § 44 определению алгебраической операции).

Важнейшим типом алгебраических систем с одной операцией являются группы. Это понятие обладает чрезвычайно широкой областью применений и служит предметом большой самостоятельной науки — теории групп. Настоящая глава может рассматриваться как введение в теорию групп — в ней будут изложены элементарные сведения о группах, знакомство с которыми необходимо каждому математику; закончится глава одной менее элементарной теоремой.

Условимся, как это принято в общей теории групп, называть рассматриваемую алгебраическую операцию *умножением* и употреблять соответствующую символику. Напомним (см. § 44), что алгебраическая операция предполагается всегда выполнимой и однозначной — для любых двух элементов a и b рассматриваемого множества произведение ab существует и является однозначно определенным элементом этого множества.

Группой называется множество G с одной алгебраической операцией, ассоциативной (хотя не обязательно коммутативной), причем для этой операции должна существовать обратная операция.

При этом, ввиду возможной некоммутативности групповой операции, выполнимость обратной операции означает следующее: для любых двух элементов a и b из G существуют в G такой одно-

значно определенный элемент x и такой однозначно определенный элемент y, что

$$ax = b, \quad ya = b.$$

Если группа G состоит из конечного числа элементов, то она называется *конечной группой*, а число элементов в ней — *порядком* группы. Если операция, определенная в группе G, коммутативна, то G называется *коммутативной* или *абелевой* группой.

Укажем простейшие следствия из определения группы. На основании рассуждений, уже проводившихся в § 44, можно утверждать, что закон ассоциативности позволяет говорить однозначным образом *о произведении любого конечного числа элементов группы*, заданных (ввиду возможной некоммутативности групповой операции) в определенном порядке.

Переходим к следствиям из существования обратной операции.

Пусть в группе G дан произвольный элемент a. Из определения группы вытекает существование в G такого однозначно определенного элемента e_a, что $ae_a = a$; этот элемент играет, следовательно, роль единицы при умножении на него элемента a справа. Если b — любой другой элемент группы G и если y есть элемент группы, удовлетворяющий равенству $ya = b$, — его существование следует из определения группы, — то мы получим:

$$b = ya = y(ae_a) = (ya)e_a = be_a.$$

Таким образом, элемент e_a играет роль правой единицы по отношению ко всем элементам группы G, а не только по отношению к исходному элементу a; поэтому мы его обозначим через e'. Из однозначности, входящей в определение обратной операции, вытекает единственность этого элемента.

Таким же путем можно доказать существование и единственность в группе G элемента e'', удовлетворяющего условию $e''a = a$ для всех a из G. На самом деле элементы e' и e'' совпадают, так как из равенств $e''e' = e''$ и $e''e' = e'$ вытекает $e'' = e'$. Этим доказано, что *во всякой группе G существует однозначно определенный элемент e, удовлетворяющий условию*

$$ae = ea = a$$

для всех a из G. Этот элемент называется *единицей* группы G и обычно обозначается символом 1.

Из определения группы вытекает, далее, существование и единственность для данного элемента a таких элементов a' и a'', что

$$aa' = 1, \quad a''a = 1.$$

В действительности элементы a' и a'' совпадают: из равенств

$$a''aa' = a''(aa') = a'' \cdot 1 = a'',$$
$$a''aa' = (a''a)a' = 1 \cdot a' = a'$$

следует $a''=a'$. Этот элемент называется *обратным* элементу a и обозначается a^{-1}, т. е.

$$aa^{-1}=a^{-1}a=1.$$

Таким образом, *всякий элемент группы обладает однозначно определенным обратным элементом.*

Из последних равенств вытекает, что обратным элементом для элемента a^{-1} служит сам элемент a. Легко видеть, далее, что обратным для произведения нескольких элементов будет произведение элементов, обратных сомножителям и притом взятых в обратном порядке:

$$(a_1a_2\ldots a_{n-1}a_n)^{-1}=a_n^{-1}a_{n-1}^{-1}\ldots a_2^{-1}a_1^{-1}.$$

Наконец, обратным элементом для единицы будет сама единица.

Проверка того, является ли группой данное множество с одной операцией, весьма облегчается тем, что в определении группы требование выполнимости обратной операции можно заменить предположением о существовании единицы и обратных элементов, причем лишь с одной стороны (например, правой) и без предположения об их единственности. Это вытекает из следующей теоремы:

Множество G с одной ассоциативной операцией будет группой, если в G существует хотя бы один элемент e, обладающий свойством

$$ae=a \quad \text{для всех } a \text{ из } G,$$

и если среди этих правых единичных элементов существует хотя бы один такой элемент e_0, что по отношению к нему всякий элемент a из G обладает хотя бы одним правым обратным элементом a^{-1}:

$$aa^{-1}=e_0.$$

Доказательство. Пусть a^{-1} — один из правых обратных элементов для a. Тогда

$$aa^{-1}=e_0=e_0e_0=e_0aa^{-1},$$

т. е. $aa^{-1}=e_0aa^{-1}$. Умножая обе части этого равенства справа на один из элементов, правых обратных для a^{-1}, мы получим $ae_0=e_0ae_0$, откуда $a=e_0a$, так как e_0 — правая единица для G. Таким образом, элемент e_0 оказывается и левой единицей для G. Если теперь e_1 есть произвольная правая единица, e_2 — произвольная левая единица, то из равенств

$$e_2e_1=e_1 \quad \text{и} \quad e_2e_1=e_2$$

следует $e_1=e_2$, т. е. любая правая единица равна любой левой. Этим доказаны существование и единственность в множестве G единичного элемента, который обозначим, как выше, через 1.

Далее,

$$a^{-1}=a^{-1}\cdot 1=a^{-1}aa^{-1}.$$

т. е. $a^{-1}=a^{-1}aa^{-1}$, где a^{-1} есть один из правых обратных элементов для a. Умножая обе части последнего равенства справа на один из правых обратных элементов для a^{-1}, мы получаем $1=a^{-1}a$, т. е. элемент a^{-1} будет служить и левым обратным элементом для a. Если теперь a_1^{-1} — произвольный правый обратный для a, a_2^{-1} — произвольный левый обратный, то из равенств

$$a_2^{-1}aa_1^{-1}=(a_2^{-1}a)\,a_1^{-1}=a_1^{-1},$$

$$a_2^{-1}aa_1^{-1}=a_2^{-1}(aa_1^{-1})=a_2^{-1}$$

следует $a_1^{-1}=a_2^{-1}$, т. е. следуют существование и единственность для всякого элемента a из G обратного элемента a^{-1}.

Теперь легко показать, что множество G будет группой. Действительно, уравнениям $ax=b$, $ya=b$ будут, как легко видеть, удовлетворять элементы

$$x=a^{-1}b, \quad y=ba^{-1}.$$

Единственность этих решений следует из того, что если, например, $ax_1=ax_2$, то, умножая обе части этого равенства слева на a^{-1}, мы получаем $x_1=x_2$. Теорема доказана.

Мы уже несколько раз встречались с понятием изоморфизма — для колец, для линейных пространств, для евклидовых пространств. Это понятие может быть определено и для групп и играет в теории групп столь же большую роль, как и в теории колец. Группы G и G' называются *изоморфными*, если между ними можно установить такое взаимно однозначное соответствие, при котором для любых элементов a, b из G и соответствующих им элементов a', b' из G' произведению ab соответствует произведение $a'b'$. Как в § 46 (для нуля и противоположного элемента в кольце), можно показать, что при изоморфном соответствии между группами G и G' единице группы G соответствует единица группы G', и если элементу a из G соответствует элемент a' из G', то элементу a^{-1} соответствует элемент a'^{-1}.

Переходя к примерам групп, отметим, что если бы операция в группе G была названа *сложением*, то единица группы называлась бы *нулем* и обозначалась символом 0, а вместо обратного элемента мы говорили бы о *противоположном элементе* и обозначали бы его через $-a$.

В качестве первого примера групп укажем, что *по сложению всякое кольцо* (*и, в частности, поле*) *является группой, притом абелевой*; это — так называемая *аддитивная группа кольца*. Это замечание сразу дает большое количество конкретных примеров групп и среди них — аддитивную группу целых чисел, аддитивную группу четных чисел, аддитивные группы рациональных чисел, действительных чисел, комплексных чисел и т. д. Заметим, что *аддитивные группы целых чисел и четных чисел изоморфны между*

собой, хотя вторая является лишь частью первой: отображение, ставящее в соответствие всякому целому числу k четное число $2k$, будет взаимно однозначным и, как легко проверить, даже изоморфным отображением первой из названных групп на вторую.

По умножению никакое кольцо не является группой, так как обратная операция — деление — не всегда выполнима. Положение не изменяется и при переходе от произвольного кольца к полю, так как в поле остается невыполнимым деление на нуль. Рассмотрим, однако, совокупность всех отличных от нуля элементов поля. Так как поле не содержит делителей нуля, т. е. произведение двух элементов, отличных от нуля, само отлично от нуля, то умножение будет для рассматриваемой совокупности алгебраической операцией, притом ассоциативной и коммутативной, причем деление уже всегда выполнимо и не выводит за пределы этой совокупности. Таким образом, ***совокупность отличных от нуля элементов любого поля является абелевой группой***; эта группа называется ***мультипликативной группой поля***. Примерами, сюда относящимися, будут мультипликативные группы рациональных чисел, действительных чисел, комплексных чисел.

Группу по умножению составляют, очевидно, все положительные действительные числа. *Эта группа изоморфна аддитивной группе всех действительных чисел*: ставя в соответствие всякому положительному числу a действительное число $\ln a$, мы получим взаимно однозначное отображение первой группы на вторую, которое будет изоморфизмом ввиду равенства

$$\ln(ab) = \ln a + \ln b.$$

Возьмем, далее, в поле комплексных чисел совокупность корней n-й степени из единицы. В § 19 было доказано, что произведение двух корней n-й степени из единицы, а также число, обратное к корню n-й степени из единицы, сами принадлежат к рассматриваемой совокупности чисел. Так как единица также принадлежит, понятно, к этой совокупности и так как умножение любых комплексных чисел ассоциативно и коммутативно, то мы получаем, что ***корни n-й степени из единицы составляют по умножению абелеву группу, притом конечную порядка n***. Таким образом, ***для любого натурального числа n существуют конечные группы порядка n***.

Группа (по умножению) корней n-й степени из единицы изоморфна аддитивной группе кольца Z_n, построенного в § 45. Действительно, если ε — первообразный корень n-й степени из единицы, то все элементы первой из названных групп имеют вид ε^k, $k = 0, 1, \ldots, n-1$. Если мы поставим в соответствие всякому числу ε^k элемент C_k кольца Z_n, т. е. класс целых чисел, дающих при делении на n остаток k, то получим изоморфное соответствие между рассматриваемыми группами: если $0 \leqslant k \leqslant n-1$, $0 \leqslant l \leqslant n-1$ и если $k+l = nq+r$, где $0 \leqslant r \leqslant n-1$, а q равно 0 или 1, то $\varepsilon^k \cdot \varepsilon^l = \varepsilon^r$ и, вместе с тем, $C_k + C_l = C_r$.

Сейчас уместно указать некоторые примеры числовых множеств, не являющихся группами. Так, множество всех целых чисел не будет группой по умножению, множество всех положительных действительных чисел не будет группой по сложению, множество всех нечетных чисел не будет группой по сложению, множество всех отрицательных действительных чисел не будет группой по умножению. Проверка всех этих утверждений не представляет затруднений.

Все рассмотренные выше числовые группы являются, конечно, абелевыми. Примерами абелевых групп, составленных не из чисел, служат линейные пространства: как вытекает из их определения (см. §§ 29, 47), *всякое линейное пространство над произвольным полем P будет абелевой группой относительно операции сложения.*

Переходим к примерам некоммутативных групп.

Множество всех матриц n-го порядка над полем P не будет группой по отношению к операции умножения, так как нарушается требование о существовании обратного элемента. Если мы ограничимся, однако, лишь невырожденными матрицами, то получим уже группу. Действительно, произведение двух невырожденных матриц будет, как мы знаем, невырожденным, единичная матрица является невырожденной, всякая невырожденная матрица обладает обратной матрицей, также невырожденной, и, наконец, закон ассоциативности, выполняясь для всех матриц, справедлив, в частности, для матриц невырожденных. Можно говорить, следовательно, о *группе невырожденных матриц n-го порядка* над полем P с умножением матриц в качестве групповой операции; эта группа некоммутативна при $n \geqslant 2$.

К весьма важным примерам конечных некоммутативных групп приводит введенное в § 3 умножение подстановок. Мы знаем, что в множестве всех подстановок n-й степени умножение будет алгебраической операцией, притом ассоциативной, хотя при $n \geqslant 3$ некоммутативной, что тождественная подстановка E служит единицей этого умножения и что для всякой подстановки существует обратная подстановка. Таким образом, *множество подстановок n-й степени составляет по умножению группу, притом конечную порядка $n!$.* Эта группа называется *симметрической группой n-й степени*; она некоммутативна при $n \geqslant 3$.

Рассмотрим теперь вместо совокупности всех подстановок n-й степени лишь совокупность четных подстановок, состоящую, как мы знаем, из $\frac{1}{2}n!$ элементов. Используя доказанную в § 3 теорему о том, что четность подстановки совпадает с четностью числа транспозиций, входящих в какое-либо разложение этой подстановки в произведение транспозиций, мы получаем, что *произведение двух четных подстановок само четно*; в самом деле, представление AB в виде произведения транспозиций мы получим, записав соответствующие разложения для A и B одно за другим. Далее, ассоциативность

умножения подстановок нам известна, четность тождественной подстановки очевидна. Наконец, четность подстановки A^{-1} при четной подстановке A следует хотя бы из того, что записи этих подстановок можно получить одну из другой переменой мест верхней и нижней строк, т. е. они содержат равное число инверсий. Таким образом, *множество четных подстановок n-й степени будет по умножению конечной группой порядка* $\frac{1}{2}n!$. Эта группа называется *знакопеременной группой n-й степени*; легко проверить, что она некоммутативна при $n \geqslant 4$, хотя будет коммутативной при $n=3$.

Симметрические и знакопеременные группы играют очень большую роль в теории конечных групп, а также в теории Галуа. Заметим, что было бы невозможно, по аналогии со знакопеременными группами, построить группу по умножению из нечетных подстановок, так как произведение двух нечетных подстановок всегда есть четная подстановка.

Большое число разнообразных примеров групп доставляют различные ветви геометрии. Укажем один простейший пример такого рода: множество всех вращений шара около его центра будет группой, притом некоммутативной, если произведением двух вращений мы назовем результат их последовательного выполнения.

§ 64. Подгруппы

Подмножество A группы G называется *подгруппой* этой группы, если оно само является группой относительно операции, определенной в группе G.

При проверке того, является ли подмножество A группы G подгруппой этой группы, достаточно проверить: 1) содержится ли в A произведение любых двух элементов из A; 2) содержит ли A вместе со всяким своим элементом и его обратный элемент. Действительно, из справедливости закона ассоциативности в группе G следует его справедливость для элементов из A, а принадлежность к A единицы группы G вытекает из 2) и 1).

Многие из групп, указанных в предшествующем параграфе, являются подгруппами других групп, также там указанных. Так, аддитивная группа четных чисел является подгруппой аддитивной группы всех целых чисел, а последняя в свою очередь есть подгруппа аддитивной группы рациональных чисел. Все эти группы, как и вообще аддитивные группы чисел, являются подгруппами аддитивной группы комплексных чисел. Мультипликативная группа положительных действительных чисел является подгруппой мультипликативной группы всех отличных от нуля действительных чисел. Знакопеременная группа n-й степени есть подгруппа симметрической группы этой же степени.

Подчеркнем, что содержащееся в определении подгруппы требование к подмножеству A группы G быть группой относительно групповой операции, определенной в группе G, является существенным. Так, мультипликативная группа положительных действительных чисел не является подгруппой аддитивной группы всех действительных чисел, хотя первое множество содержится, как подмножество, во втором.

Если в группе G взяты подгруппы A и B, то их пересечение $A \cap B$, т. е. совокупность элементов, лежащих и в A, и в B, также будет подгруппой группы G.

Действительно, если в пересечении $A \cap B$ содержатся элементы x и y, то они лежат в подгруппе A, а поэтому к A принадлежит и произведение xy, и обратный элемент x^{-1}. По тем же соображениям элементы xy и x^{-1} принадлежат и к подгруппе B, а поэтому они входят и в $A \cap B$.

Полученный результат справедлив, как легко видеть, не только для двух подгрупп, но и для любого числа подгрупп, конечного или даже бесконечного.

Подмножество группы G, состоящее из одного элемента 1, будет, очевидно, подгруппой этой группы; эта подгруппа, содержащаяся в любой другой подгруппе группы G, называется *единичной подгруппой* группы G. С другой стороны, сама группа G является одной из своих подгрупп.

Интересным примером подгрупп служат так называемые циклические подгруппы. Введем сначала понятие степени элемента a группы G. Если n — любое натуральное число, то произведение n элементов, равных элементу a, называется *n-й степенью* элемента a и обозначается через a^n. *Отрицательные степени* элемента a можно определить или как элементы группы G, обратные положительным степеням этого элемента, или же как произведения нескольких множителей, равных элементу a^{-1}. В действительности эти определения совпадают,

$$(a^n)^{-1} = (a^{-1})^n, \quad n > 0. \tag{1}$$

Для доказательства достаточно взять произведение $2n$ множителей, из которых первые n равны a, а остальные равны a^{-1}, и произвести все сокращения. Элемент, равный как левой, так и правой части равенства (1), будет обозначаться через a^{-n}. Условимся, наконец, под *нулевой степенью* a^0 элемента a понимать элемент 1.

Заметим, что если операция в группе G называется сложением, то вместо степеней элемента a следует говорить о *кратных* этого элемента и записывать их через ka.

Без труда проверяется, что в любой группе G для степеней любого элемента a при любых показателях m и n, положительных,

отрицательных или нулевых, имеют место равенства

$$a^n \cdot a^m = a^m \cdot a^n = a^{n+m}, \tag{2}$$

$$(a^n)^m = a^{nm}. \tag{3}$$

Обозначим через $\{a\}$ подмножество группы G, составленное из всех степеней элемента a; в него входит и сам элемент a, являющийся своей первой степенью. *Подмножество $\{a\}$ будет подгруппой группы G*: произведение элементов из $\{a\}$ лежит в $\{a\}$ ввиду (2), в $\{a\}$ входит элемент 1, равный a^0, и, наконец, $\{a\}$ вместе со всяким своим элементом содержит и его обратный элемент, так как из (3) следует равенство

$$(a^n)^{-1} = a^{-n}.$$

Подгруппа $\{a\}$ называется *циклической подгруппой группы G, порожденной элементом a*. Как показывает равенство (2), она всегда коммутативна, даже если сама группа G и некоммутативна.

Заметим, что нигде выше не утверждалось, что все степени элемента a являются различными элементами группы. Если это действительно так, то a называется *элементом бесконечного порядка*. Пусть, однако, среди степеней элемента a имеются равные, например, $a^k = a^l$ при $k \neq l$; это всегда имеет место в случае конечных групп, но может случиться и в бесконечной группе. Если $k > l$, то

$$a^{k-l} = 1,$$

т. е. существуют положительные степени элемента a, равные единице. Пусть n есть наименьшая положительная степень элемента a, равная единице, т. е.

1) $a^n = 1, \quad n > 0,$

2) если $a^k = 1, \quad k > 0, \quad$ то $\quad k \geqslant n.$

В этом случае говорят, что a есть *элемент конечного порядка*, а именно *порядка n*.

Если элемент a имеет конечный порядок n, то все элементы

$$1, a, a^2, \ldots, a^{n-1} \tag{4}$$

будут, как легко видеть, различными. *Всякая другая степень элемента a, положительная или отрицательная, равна одному из элементов* (4). Действительно, если k — любое целое число, то, деля k на n, получим

$$k = nq + r, \quad 0 \leqslant r \leqslant n,$$

а поэтому, ввиду (2) и (3),

$$a^k = (a^n)^q \cdot a^r = a^r. \tag{5}$$

Отсюда следует, что *если элемент a имеет конечный порядок n и $a^k=1$, то k должно нацело делиться на n.* С другой стороны, так как

$$-1=n(-1)+(n-1),$$

то *для элемента a конечного порядка n*

$$a^{-1}=a^{n-1}.$$

Так как система (4) содержит n элементов, то из полученных выше результатов вытекает, что *для элемента a, имеющего конечный порядок, его порядок n совпадает с порядком* (*т. е. с числом элементов*) *циклической подгруппы* $\{a\}$.

Заметим, наконец, что всякая группа обладает одним-единственным элементом первого порядка — это будет элемент 1. Циклическая подгруппа $\{1\}$ совпадает, очевидно, с единичной подгруппой.

Циклические группы. Группа G называется *циклической* группой, если она состоит из степеней одного из своих элементов a, т. е. совпадает с одной из своих циклических подгрупп $\{a\}$; элемент a называется в этом случае *образующим элементом* группы G. Всякая циклическая группа, очевидно, абелева.

Примером бесконечной циклической группы служит аддитивная группа целых чисел — всякое целое число кратно числу 1, т. е. это число служит образующим элементом рассматриваемой группы; в качестве образующего элемента можно было бы взять также число -1.

Примером конечной циклической группы порядка n служит мультипликативная группа корней n-й степени из единицы — в § 19 показано, что все эти корни являются степенями одного из них, а именно первообразного корня.

Следующая теорема показывает, что этими примерами исчерпываются по существу все циклические группы:

Все бесконечные циклические группы изоморфны между собой; изоморфны между собой также все конечные циклические группы данного порядка n.

Действительно, бесконечная циклическая группа с образующим элементом a отображается взаимно однозначно на аддитивную группу целых чисел, если всякому элементу a^k этой группы ставится в соответствие число k; это отображение будет изоморфным, так как, по (2), при перемножении степеней элемента a показатели складываются. Если же дана конечная циклическая группа G порядка n с образующим элементом a, то обозначим через ε первообразный корень n-й степени из единицы и сопоставим всякому элементу a^k группы G, $0\leqslant k<n$, число ε^k. Это будет взаимно однозначное отображение группы G на мультипликативную группу корней n-й степени из 1, изоморфность которого следует из (2) и (5).

Эта теорема позволяет говорить просто о *бесконечной циклической группе* или о *циклической группе порядка n*.

Докажем, далее, следующую теорему:

Всякая подгруппа циклической группы сама циклическая.

В самом деле, пусть $G=\{a\}$ есть циклическая группа с образующим элементом a, бесконечная или конечная, и пусть A будет подгруппа группы G. Можно считать, что A отлична от единичной подгруппы, так как иначе доказывать было бы нечего. Предположим, что a^k есть наименьшая положительная степень элемента a, содержащаяся в A; такая степень существует, так как если в A содержится отличный от 1 элемент a^{-s}, $s>0$, то содержится и обратный ему элемент a^s. Допустим, что в A содержится также элемент a^l, $l\neq 0$, причем l не делится на k. Тогда, если d, $d>0$, есть наибольший общий делитель чисел k и l, то существуют такие целые числа u и v, что

$$ku+lv=d,$$

а поэтому в подгруппе A должен содержаться элемент

$$(a^k)^u\cdot(a^l)^v=a^d,$$

но так как при наших предположениях $d<k$, то мы приходим в противоречие с выбором элемента a^k. Этим доказано, что $A=\{a^k\}$.

Разложения группы по подгруппе. Если в группе G взяты подмножества M и N, то под *произведением* MN *этих подмножеств* понимается совокупность элементов группы G, которые хотя бы одним способом представимы в виде произведения некоторого элемента из M на некоторый элемент из N. Из ассоциативности групповой операции следует *ассоциативность умножения подмножеств группы*,

$$(MN)P=M(NP).$$

Одно из множеств M, N может состоять, понятно, лишь из одного элемента a. В этом случае мы получаем *произведение* aN *элемента на множество или произведение* Ma *множества на элемент*.

Пусть в группе G дана произвольная подгруппа A. Если x — любой элемент из G, то произведение xA называется *левым смежным классом группы* G *по подгруппе* A, *порождаемым элементом* x. Понятно, что *элемент* x *содержится в смежном классе* xA, так как подгруппа A содержит единицу, но $x\cdot 1=x$.

Всякий левый смежный класс порождается любым из своих элементов, т. е. если элемент y содержится в смежном классе xA, то

$$yA=xA. \qquad (6)$$

Действительно, y можно представить в виде

$$y=xa,$$

где a — элемент подгруппы A. Поэтому для любых элементов a' и a'' из A будет

$$ya' = x(aa'),$$
$$xa'' = y(a^{-1}a''),$$

чем и доказывается равенство (6).

Отсюда следует, что *два любых левых смежных класса группы G по подгруппе A или совпадают, или же не имеют ни одного общего элемента*. Действительно, если смежные классы xA и yA содержат общий элемент z, то

$$xA = zA = yA.$$

Таким образом, вся группа G распадается на непересекающиеся левые смежные классы по подгруппе A. Это разложение называется *левосторонним разложением группы G по подгруппе A*.

Заметим, что одним из левых смежных классов этого разложения будет сама подгруппа A; этот смежный класс порождается элементом 1 или, вообще, любым элементом a из A, так как

$$aA = A.$$

Понятно, что, называя *правым смежным классом группы G по подгруппе A, порождаемым элементом x*, произведение Ax, мы аналогичным путем получили бы *правостороннее разложение группы G по подгруппе A*. Для абелевой группы оба ее разложения по любой подгруппе, левостороннее и правостороннее, будут, понятно, совпадать, т. е. можно говорить просто о *разложении группы по подгруппе*.

Так, разложение аддитивной группы целых чисел по подгруппе чисел, кратных числу k, состоит из k различных смежных классов, порождаемых соответственно числами $0, 1, 2, \ldots, k-1$. При этом в классе, порождаемом числом l, $0 \leqslant l \leqslant k-1$, собраны все те числа, которые при делении на k дают остаток l.

В некоммутативном случае разложения группы по некоторой подгруппе могут оказаться различными.

Рассмотрим, например, симметрическую группу 3-й степени S_3, причем, в соответствии с § 3, будем записывать ее элементы через циклы. В качестве подгруппы A возьмем циклическую подгруппу элемента (12); она состоит из тождественной подстановки и самой подстановки (12). Другими левыми смежными классами будут: класс $(13)\cdot A$, состоящий из подстановок (13) и (132), и класс $(23)\cdot A$, состоящий из подстановок (23) и (123). С другой стороны, правыми смежными классами группы S_3 по подгруппе A будут: сама подгруппа A, класс $A\cdot(13)$, состоящий из подстановок (13) и (123), и класс $A\cdot(23)$, состоящий из подстановок (23) и (132). Мы видим, что правостороннее разложение отличается в рассматриваемом случае от левостороннего.

Для случая конечных групп существование разложений группы по подгруппе приводит к следующей важной теореме:

Теорема Лагранжа. *Во всякой конечной группе порядок любой подгруппы является делителем порядка самой группы.*

В самом деле, пусть в конечной группе G порядка n дана подгруппа A порядка k. Рассмотрим левостороннее разложение группы G по подгруппе A. Пусть оно состоит из j классов; число j называется *индексом* подгруппы A в группе G. Каждый левый класс xA состоит ровно из k элементов, так как если

$$xa_1 = xa_2,$$

где a_1 и a_2 — элементы из A, то $a_1 = a_2$. Таким образом,

$$n = kj, \tag{7}$$

что и требовалось доказать.

Так как порядок элемента совпадает с порядком его циклической подгруппы, то из теоремы Лагранжа следует, что *порядок всякого элемента конечной группы является делителем порядка группы.*

Из теоремы Лагранжа следует также, что *всякая конечная группа, порядок которой есть простое число, будет циклической.* Действительно, эта группа должна совпадать с циклической подгруппой, порожденной любым ее элементом, отличным от единицы. Отсюда вытекает, ввиду полученного выше описания циклических групп, что *для всякого простого числа p существует единственная, с точностью до изоморфизма, конечная группа порядка p.*

§ 65. Нормальные делители, фактор-группы, гомоморфизмы

Подгруппа A группы G называется *нормальным делителем* этой группы (или *инвариантной подгруппой*), если левостороннее разложение группы G по подгруппе A совпадает с правосторонним.

Таким образом, все подгруппы абелевой группы являются в ней нормальными делителями. С другой стороны, во всякой группе G и единичная подгруппа, и сама эта группа будут нормальными делителями: оба разложения группы G по единичной подгруппе совпадают с разложением группы на отдельные элементы, оба разложения группы G по самой этой группе состоят из одного класса G.

Укажем более интересные примеры нормальных делителей в некоммутативных группах. В симметрической группе 3-й степени S_3 циклическая подгруппа элемента (123), состоящая из тождественной подстановки и подстановок (123) и (132), будет нормальным делителем: в обоих разложениях группы S_3 по этой подгруппе второй смежный класс состоит из подстановок (12), (13) и (23).

Вообще в симметричной группе n-й степени S_n знакопеременная группа n-й степени A_n будет нормальным делителем. Действительно, группа A_n имеет порядок $\frac{1}{2}n!$, поэтому всякий смежный класс

группы S_n по подгруппе A_n должен состоять из стольких же элементов и, следовательно, такой класс имеется еще только один, а именно совокупность нечетных подстановок.

В мультипликативной группе невырожденных квадратных матриц порядка n с элементами из поля P те матрицы, определитель которых равен 1, составляют, очевидно, подгруппу. Это будет даже нормальный делитель, так как смежным классом по этой подгруппе, одновременно левым и правым, порождаемым матрицей M, является класс всех матриц, определитель которых равен определителю матрицы M — достаточно вспомнить, что при умножении матриц их определители перемножаются.

Определению нормального делителя, приведенному выше, можно придать такую форму:

Подгруппа A группы G называется нормальным делителем этой группы, если для всякого элемента x из G

$$xA = Ax, \tag{1}$$

т. е. для всякого элемента x из G и элемента a из A можно подобрать в A такие элементы a' и a'', что

$$xa = a'x, \quad ax = xa''. \tag{2}$$

Можно указать и другие определения нормального делителя, равносильные исходному. Так, назовем элементы a и b группы G *сопряженными*, если в G существует хотя бы один такой элемент x, что

$$b = x^{-1}ax, \tag{3}$$

т. е., как говорят, элемент b получается из элемента a *трансформированием* элементом x. Из (3) следует, очевидно, равенство

$$a = xbx^{-1} = (x^{-1})^{-1}bx^{-1}.$$

Подгруппа A группы G тогда и только тогда будет нормальным делителем в G, если вместе со всяким своим элементом a она содержит и все элементы, сопряженные с ним в G.

Действительно, если A — нормальный делитель в G, то, по (2), для выбранного нами элемента a из A и любого элемента x из G можно подобрать в A такой элемент a'', что

$$ax = xa''.$$

Отсюда

$$x^{-1}ax = a'',$$

т. е. всякий элемент, сопряженный с a, содержится в A. Обратно, если подгруппа A содержит вместе со всяким своим элементом a и все элементы, ему сопряженные, то в A содержится, в частности, элемент

$$x^{-1}ax = a'',$$

откуда следует второе из равенств (2). По той же причине в A содержится и элемент

$$(x^{-1})^{-1}ax^{-1} = xax^{-1} = a',$$

откуда следует и первое из равенств (2).

Пользуясь этим результатом, легко доказать, что *пересечение любых нормальных делителей группы G само будет нормальным делителем этой группы.* В самом деле, если A и B — нормальные делители группы G, то, как показано в предыдущем параграфе, пересечение $A \cap B$ будет подгруппой группы G. Пусть c — любой элемент из $A \cap B$, x — любой элемент группы G. Тогда элемент $x^{-1}cx$ должен содержаться и в A, и в B, так как оба этих нормальных делителя содержат элемент c. Отсюда следует, что элемент $x^{-1}cx$ входит в пересечение $A \cap B$.

Фактор-группа. Значение понятия нормального делителя основано на том, что из смежных классов по нормальному делителю — ввиду (1) левые и правые смежные классы можно в этом случае не различать — некоторым весьма естественным способом может быть построена новая группа.

Заметим сначала, что если A — произвольная подгруппа группы G, то

$$AA = A, \tag{4}$$

так как произведение любых двух элементов из подгруппы A принадлежит к A и, вместе с тем, умножая все элементы из A на единицу, мы уже получим всю подгруппу A.

Пусть A будет теперь нормальным делителем группы G. *В этом случае произведение любых двух смежных классов G по A* (в смысле умножения подмножеств группы G) *само будет смежным классом по A.* Действительно, используя ассоциативность умножения подмножеств группы, равенство (4) и равенство

$$yA = Ay$$

(ср. (1)), мы для любых элементов x и y группы G получим:

$$xA \cdot yA = xyAA = xyA. \tag{5}$$

Равенство (5) показывает, что для того, чтобы найти произведение двух данных смежных классов группы G по нормальному делителю A, следует произвольным образом выбрать в этих смежных классах по одному п р е д с т а в и т е л ю — напомним, что всякий смежный класс порождается любым из своих элементов — и взять тот смежный класс, в котором лежит произведение этих представителей.

Таким образом, в множестве всех смежных классов группы G по нормальному делителю A определена операция умножения. Покажем, что *при этом выполняются все требования, входящие в определение группы.* В самом деле, ассоциативность умножения

смежных классов следует из ассоциативности умножения подмножеств группы. Роль единицы играет сам нормальный делитель A, являющийся одним из смежных классов разложения G по A: именно, ввиду (4) и (1) для любого x из G будет

$$xA \cdot A = xA, \quad A \cdot xA = xAA = xA.$$

Наконец, для смежного класса xA обратным будет смежный класс $x^{-1}A$, так как

$$xA \cdot x^{-1}A = 1 \cdot A = A.$$

Построенная нами группа называется *фактор-группой* группы G по нормальному делителю A и обозначается через G/A.

Мы видим, что со всякой группой связывается целый набор новых групп — ее фактор-групп по различным нормальным делителям. При этом фактор-группа группы G по единичной подгруппе будет, понятно, изоморфной с самой группой G.

Всякая фактор-группа G/A абелевой группы G сама является абелевой, так как из $xy = yx$ следует

$$xA \cdot yA = xyA = yxA = yA \cdot xA.$$

Всякая фактор-группа G/A циклической группы G сама циклическая, так как если G порождается элементом g, $G = \{g\}$, и если дан произвольный смежный класс xA, то существует такое целое число k, что

$$x = g^k$$

и поэтому

$$xA = (gA)^k.$$

Порядок любой фактор-группы G/A конечной группы G является делителем порядка самой этой группы. Действительно, порядок фактор-группы G/A равен индексу нормального делителя A в группе G, а поэтому можно воспользоваться равенством (7) из предшествующего параграфа.

Приведем некоторые примеры фактор-групп. Так как в аддитивной группе целых чисел подгруппа чисел, кратных натуральному числу k, имеет, как показано в предшествующем параграфе, индекс k, то фактор-группа нашей группы по этой подгруппе будет конечной группой порядка k, притом циклической, так как сама рассматриваемая группа циклическая.

Фактор-группа симметричной группы n-й степени S_n по знакопеременной группе n-й степени A_n будет группой 2-го порядка, причем, ввиду простоты числа 2, циклической группой (см. конец предшествующего параграфа).

Выше приведено описание смежных классов мультипликативной группы невырожденных матриц порядка n с элементами из поля P по нормальному делителю, составленному из матриц, определитель

которых равен 1. Из этого описания следует, что соответствующая фактор-группа изоморфна мультипликативной группе отличных от нуля чисел поля P.

Гомоморфизмы. Понятия нормального делителя и фактор-группы тесно связаны со следующим обобщением понятия изоморфизма.

Отображение φ группы G на группу G', ставящее в соответствие всякому элементу a из G однозначно определенный элемент $a' = a\varphi$ из G', называется *гомоморфным отображением* G на G' (или просто *гомоморфизмом*), если всякий элемент a' из G' служит при этом отображении образом некоторого элемента a из G, $a' = a\varphi$, и если для любых элементов a, b группы G

$$(ab)\varphi = a\varphi \cdot b\varphi.$$

Очевидно, что, потребовав дополнительно взаимную однозначность отображения φ, мы получили бы уже известное нам определение изоморфизма.

Если φ — гомоморфизм группы G на группу G' и 1 и a — соответственно единица и произвольный элемент группы G, $1'$ — единица группы G', то

$$1\varphi = 1',$$
$$(a^{-1})\varphi = (a\varphi)^{-1}.$$

Действительно, если $1\varphi = e'$ и x' — произвольный элемент группы G', то в G существует такой элемент x, что $x\varphi = x'$. Отсюда

$$x' = x\varphi = (x \cdot 1)\varphi = x\varphi \cdot 1\varphi = x' \cdot e'.$$

Аналогично

$$x' = e'x'$$

и, следовательно, $e' = 1'$.

С другой стороны, если $(a^{-1})\varphi = b'$, то

$$1' = 1\varphi = (aa^{-1})\varphi = a\varphi \cdot (a^{-1})\varphi = a\varphi \cdot b'$$

и, аналогично,

$$1' = b' \cdot a\varphi,$$

откуда $b' = (a\varphi)^{-1}$.

Назовем *ядром* гомоморфизма φ группы G на группу G' совокупность тех элементов группы G, которые отображаются при φ в единицу $1'$ группы G'.

Ядро всякого гомоморфизма φ группы G является нормальным делителем группы G.

Действительно, если элементы a, b группы G входят в ядро гомоморфизма φ, т. е.

$$a\varphi = b\varphi = 1',$$

то

$$(ab)\varphi = a\varphi \cdot b\varphi = 1' \cdot 1' = 1',$$

т. е. и произведение ab содержится в ядре гомоморфизма φ. С другой стороны, если $a\varphi = 1'$, то

$$(a^{-1})\varphi = (a\varphi)^{-1} = 1'^{-1} = 1',$$

т. е. и a^{-1} входит в ядро гомоморфизма φ. Наконец, если $a\varphi = 1'$, а x — произвольный элемент группы G, то

$$(x^{-1}ax)\varphi = (x^{-1})\varphi \cdot a\varphi \cdot x\varphi = (x\varphi)^{-1} \cdot 1' \cdot x\varphi = 1'.$$

Ядро рассматриваемого гомоморфизма оказалось подгруппой группы G, содержащей вместе со всяким своим элементом и все элементы, с ним сопряженные; оно будет, следовательно, нормальным делителем.

Пусть теперь A — произвольный нормальный делитель группы G. Ставя в соответствие всякому элементу x группы G тот смежный класс xA по нормальному делителю A, в котором этот элемент лежит, мы получим отображение группы G на всю фактор-группу G/A. Из определения умножения в группе G/A (см. (5)) следует, что это отображение будет гомоморфным.

Полученный гомоморфизм называется *естественным гомоморфизмом* группы G на фактор-группу G/A. Ядром этого гомоморфизма служит, очевидно, сам нормальный делитель A.

Отсюда следует, что *нормальные делители группы G и только они служат ядрами гомоморфизмов этой группы.* Этот результат можно рассматривать как еще одно определение нормального делителя.

Оказывается, что все группы, на которые группа G может гомоморфно отображаться, по существу исчерпываются фактор-группами этой группы, а все гомоморфизмы группы G — ее естественными гомоморфизмами на свои фактор-группы. Точнее, справедлива следующая

Теорема о гомоморфизмах. *Пусть дан гомоморфизм φ группы G на группу G' и пусть A — ядро этого гомоморфизма. Тогда группа G' изоморфна фактор-группе G/A, причем существует такое изоморфное отображение σ первой из этих групп на вторую, что результат последовательного выполнения отображений φ и σ совпадает с естественным гомоморфизмом группы G на фактор-группу G/A.*

В самом деле, пусть x' будет произвольный элемент группы G', а x — такой элемент группы G, что $x\varphi = x'$. Так как для любого элемента a из ядра A гомоморфизма φ имеет место равенство $a\varphi = 1'$, то

$$(xa)\varphi = x\varphi \cdot a\varphi = x' \cdot 1' = x',$$

т. е. все элементы смежного класса xA отображаются при φ в элемент x'.

С другой стороны, если z — любой такой элемент группы G, что $z\varphi = x'$, то

$$(x^{-1}z)\varphi = x^{-1}\varphi \cdot z\varphi = (x\varphi)^{-1} \cdot z\varphi = x'^{-1} \cdot x' = 1',$$

т. е. $x^{-1}z$ содержится в ядре A гомоморфизма φ. Если мы положим $x^{-1}z = a$, то $z = xa$, т. е. элемент z содержится в смежном классе xA. Таким образом, собирая все те элементы группы G, которые при гомоморфизме φ отображаются в фиксированный элемент x' группы G', мы получаем точно смежный класс xA.

Соответствие σ, относящее каждому элементу x' из G' тот смежный класс группы G по нормальному делителю A, который состоит из всех элементов группы G, имеющих x' своим образом при φ, будет взаимно однозначным отображением группы G' на группу G/A. Это отображение σ будет изоморфизмом, так как если

$$x'\sigma = xA, \quad y'\sigma = yA,$$

т. е.

$$x\varphi = x', \quad y\varphi = y',$$

то

$$(xy)\varphi = x\varphi \cdot y\varphi = x'y',$$

а поэтому

$$(x'y')\sigma = xyA = xA \cdot yA = x'\sigma \cdot y'\sigma.$$

Наконец, если x — произвольный элемент из G и $x\varphi = x'$, то

$$(x\varphi)\sigma = x'\sigma = xA,$$

т. е. последовательное выполнение гомоморфизма φ и изоморфизма σ на самом деле отображает элемент x в порождаемый им смежный класс xA. Теорема доказана.

§ 66. Прямые суммы абелевых групп

Мы хотим закончить главу одной теоретико-групповой теоремой, более глубокой, чем те элементарные свойства групп, которые излагались выше. Именно, опираясь на уже известное нам из § 64 описание циклических групп, мы получим в следующем параграфе *полное описание конечных абелевых групп*.

Как принято в теории абелевых групп, для групповой операции будет использоваться аддитивная запись: мы будем говорить о *сумме* $a+b$ элементов a и b группы, о *нулевой подгруппе* 0, о *кратных* ka некоторого элемента a и т. д.

В этом параграфе мы изучим одну конструкцию, которую будем излагать применительно к абелевым группам, хотя ее можно было бы вводить сразу для любых (т. е. не обязательно коммутативных) групп. Эта конструкция подсказывается следующими примерами. Плоскость, рассматриваемая как двумерное действительное линейное пространство, является абелевой группой относительно сложения

векторов. Любая прямая в этой плоскости, проходящая через начало координат, будет подгруппой указанной группы. Если A_1 и A_2 — две различные такие прямые, то, как известно, всякий вектор на плоскости, выходящий из начала координат, однозначно представляется в виде суммы своих проекций на прямые A_1 и A_2. Аналогично, всякий вектор трехмерного линейного пространства однозначно записывается в виде суммы трех векторов, принадлежащих к трем заданным прямым A_1, A_2 и A_3, если только эти прямые не лежат в одной плоскости.

Абелева группа G называется *прямой суммой* своих подгрупп $A_1, A_2, \ldots, A_k$,

$$G = A_1 + A_2 + \ldots + A_k, \tag{1}$$

если всякий элемент x группы G записывается, притом единственным способом, в виде суммы элементов $a_1, a_2, \ldots, a_k$, взятых соответственно в подгруппах $A_1, A_2, \ldots, A_k$,

$$x = a_1 + a_2 + \ldots + a_k. \tag{2}$$

Запись (1) называется *прямым разложением* группы G, подгруппы A_i, $i = 1, 2, \ldots, k$, — *прямыми слагаемыми* этого разложения, а элемент a_i из (2) — *компонентой* элемента x в прямом слагаемом A_i разложения (1), $i = 1, 2, \ldots, k$.

Если дано прямое разложение (1) *группы G и если прямые слагаемые A_i этого разложения, все или некоторые, сами разложены в прямую сумму,*

$$A_i = A_{i1} + A_{i2} + \ldots + A_{ik_i}, \quad k_i \geqslant 1, \tag{3}$$

то группа G будет прямой суммой всех своих подгрупп

$$A_{ij}, \quad j = 1, 2, \ldots, k_i, \quad i = 1, 2, \ldots, k.$$

Действительно, для произвольного элемента x группы G существует запись (2) относительно прямого разложения (1), а для каждой компоненты a_i, $i = 1, 2, \ldots, k$, — запись

$$a_i = a_{i1} + a_{i2} + \ldots + a_{ik_i} \tag{4}$$

относительно прямого разложения (3) группы A_i. Ясно, что x будет суммой всех элементов a_{ij}, $j = 1, 2, \ldots, k_i$, $i = 1, 2, \ldots, k$. Единственность этой записи вытекает из того, что, беря любую запись элемента x в виде суммы элементов, взятых по одному в подгруппах A_{ij}, и складывая слагаемые, принадлежащие к одной и той же подгруппе A_i, $i = 1, 2, \ldots, k$, мы должны получить как раз равенство (2); с другой стороны, каждый элемент a_i обладает лишь одной записью вида (4).

Определению прямой суммы можно придать иную форму. Введем сначала еще одно понятие. Если в абелевой группе G даны

некоторые подгруппы $B_1, B_2, \ldots, B_l$, то обозначим через $\{B_1, B_2, \ldots, B_l\}$ совокупность элементов y группы G, которые хотя бы одним способом могут быть записаны в виде суммы элементов $b_1, b_2, \ldots, b_l$, взятых соответственно в подгруппах $B_1, B_2, \ldots, B_l$,

$$y = b_1 + b_2 + \ldots + b_l. \qquad (5)$$

Множество $\{B_1, B_2, \ldots, B_l\}$ *будет подгруппой группы* G. Говорят, что эта подгруппа *порождена* подгруппами $B_1, B_2, \ldots, B_l$.

Для доказательства возьмем в $\{B_1, B_2, \ldots, B_l\}$ элемент y с записью (5), а также элемент y', обладающий аналогичной записью

$$y' = b'_1 + b'_2 + \ldots + b'_l,$$

где b'_i — элемент из B_i, $i = 1, 2, \ldots, l$. Тогда

$$y + y' = (b_1 + b'_1) + (b_2 + b'_2) + \ldots + (b_l + b'_l),$$
$$-y = (-b_1) + (-b_2) + \ldots + (-b_l);$$

т. е. элементы $y + y'$ и $-y$ также обладают хотя бы одной записью вида (5) и, следовательно, принадлежат к множеству $\{B_1, B_2, \ldots, B_l\}$, что и требовалось доказать.

Подгруппа $\{B_1, B_2, \ldots, B_l\}$ *содержит каждую из подгрупп* B_i, $i = 1, 2, \ldots, l$. Действительно, всякая подгруппа группы G содержит нуль этой группы, а поэтому, беря, например, в подгруппе B_1 любой элемент b_1, а в подгруппах $B_2, \ldots, B_l$ — элемент 0, мы получим для элемента b_1 следующую запись вида (5):

$$b_1 = b_1 + 0 + \ldots + 0.$$

Абелева группа G *тогда и только тогда будет прямой суммой своих подгрупп* $A_1, A_2, \ldots, A_k$, *если она порождается этими подгруппами,*

$$G = \{A_1, A_2, \ldots, A_k\}, \qquad (6)$$

и если пересечение каждой подгруппы A_i, $i = 2, \ldots, k$, *с подгруппой, порожденной всеми предшествующими подгруппами* $A_1, A_2, \ldots, A_{i-1}$, *содержит только нуль,*

$$\{A_1, A_2, \ldots, A_{i-1}\} \cap A_i = 0, \quad i = 2, \ldots, k. \qquad (7)$$

Действительно, если группа G обладает прямым разложением (1), то для всякого элемента x из G существует запись (2), а поэтому имеет место равенство (6). Справедливость равенств (7) вытекает из единственности записи (2) для любого элемента x: если бы для некоторого i пересечение $\{A_1, A_2, \ldots, A_{i-1}\} \cap A_i$ содержало ненулевой элемент x, то, с одной стороны, x можно записать как элемент a_i из A_i, т. е. $x = a_i$, и поэтому

$$x = 0 + \ldots + 0 + a_i + 0 + \ldots + 0; \qquad (8)$$

с другой стороны, x, как элемент из подгруппы $\{A_1, A_2, \ldots, A_{i-1}\}$ обладает записью вида

$$x = a_1 + a_2 + \ldots + a_{i-1},$$

т. е.

$$x = a_1 + a_2 + \ldots + a_{i-1} + 0 + \ldots + 0. \qquad (9)$$

(8) и (9) будут, очевидно, двумя разными записями вида (2) для элемента x.

Обратно, пусть выполняются равенства (6) и (7). Из (6) следует, что любой элемент x группы G обладает хотя бы одной записью вида (2). Пусть, однако, для некоторого элемента x существуют две различные записи вида (2),

$$x = a_1 + a_2 + \ldots + a_k = a'_1 + a'_2 + \ldots + a'_k. \qquad (10)$$

Тогда можно найти такое i, $i \leqslant k$, что

$$a_k = a'_k, \quad a_{k-1} = a'_{k-1}, \quad \ldots, \quad a_{i+1} = a'_{i+1}, \qquad (11)$$

но

$$a_i \neq a'_i,$$

т. е.

$$a_i - a'_i \neq 0. \qquad (12)$$

Из (10) и (11) следует, однако, равенство

$$a_i - a'_i = (a'_1 - a_1) + (a'_2 - a_2) + \ldots + (a'_{i-1} - a_{i-1}),$$

противоречащее, ввиду (12), равенству (7). Теорема доказана.

На понятие прямой суммы можно посмотреть с совсем иной стороны. Пусть дано k произвольных абелевых групп $A_1, A_2, \ldots, A_k$, среди которых могут быть и изоморфные. Обозначим через G совокупность всевозможных систем вида

$$(a_1, a_2, \ldots, a_k), \qquad (13)$$

составленных из элементов, взятых по одному в каждой из групп $A_1, A_2, \ldots, A_k$. Множество G станет абелевой группой, если *сложение* систем вида (13) будет определено правилом:

$$(a_1, a_2, \ldots, a_k) + (a'_1, a'_2, \ldots, a'_k) =$$
$$= (a_1 + a'_1, a_2 + a'_2, \ldots, a_k + a'_k), \qquad (14)$$

т. е. складываются элементы в каждой из заданных групп $A_1, A_2, \ldots, A_k$ отдельно. Действительно, ассоциативность и коммутативность этого сложения вытекают из справедливости этих свойств в каждой из заданных групп; роль нуля играет система

$$(0_1, 0_2, \ldots, 0_k),$$

где через 0_i обозначен нулевой элемент группы A_i, $i=1, 2, \ldots, k$; противоположной для системы (13) будет система

$$(-a_1, -a_2, \ldots, -a_k).$$

Построенная абелева группа G называется *прямой суммой* групп $A_1, A_2, \ldots, A_k$ и записывается, как и выше, через

$$G=A_1+A_2+\ldots+A_k.$$

Оправданием для этого названия служит то, что *группа G, являющаяся прямой суммой групп $A_1, A_2, \ldots, A_k$ в только что определенном смысле, может быть разложена в прямую сумму своих подгрупп $A_1', A_2', \ldots, A_k'$, соответственно изоморфных группам $A_1, A_2, \ldots, A_k$.*

Именно, обозначим через A_i', $i=1, 2, \ldots, k$, совокупность тех элементов группы G, т. е. систем вида (13), у которых на i-м месте стоит произвольный элемент a_i из группы A_i, а все остальные места заняты нулями соответствующих групп; это будут, следовательно, системы вида

$$(0_1, \ldots, 0_{i-1}, a_i, 0_{i+1}, \ldots, 0_k). \qquad (15$$

Определение сложения (14) показывает, что множество A_i' будет подгруппой группы G; изоморфизм этой подгруппы с группой A_i мы получим, сопоставляя каждой системе (15) элемент a_i группы A_i.

Остается доказать, что группа G является прямой суммой подгрупп $A_1', A_2', \ldots, A_k'$. Действительно, любой элемент (13) группы G можно представить в виде суммы элементов из указанных подгрупп:

$$(a_1, a_2, \ldots, a_k)=(a_1, 0_2, \ldots, 0_k)+$$
$$+(0_1, a_2, 0_3, \ldots, 0_k)+\ldots+(0_1, 0_2, \ldots, 0_{k-1}, a_k).$$

Единственность этого представления вытекает из того, что различные системы вида (13) являются различными элементами группы G.

Если даны две системы абелевых групп, $A_1, A_2, \ldots, A_k$ и $B_1, B_2, \ldots, B_k$, причем группы A_i и B_i изоморфны, $i=1, 2, \ldots, k$, то группы

$$G=A_1+A_2+\ldots+A_k$$

и

$$H=B_1+B_2+\ldots+B_k$$

также будут изоморфными.

Действительно, если для $i=1, 2, \ldots, k$ между группами A_i и B_i установлен изоморфизм φ_i, сопоставляющий каждому элементу a_i из A_i элемент $a_i\varphi_i$ из B_i, то отображение φ, относящее

всякому элементу $(a_1, a_2, \ldots, a_k)$ группы G элемент группы H, определяемый равенством

$$(a_1, a_2, \ldots, a_k)\varphi = (a_1\varphi_1, a_2\varphi_2, \ldots, a_k\varphi_k),$$

будет, очевидно, изоморфным отображением группы G на группу H.

Если даны конечные абелевы группы $A_1, A_2, \ldots, A_k$, имеющие соответственно порядки $n_1, n_2, \ldots, n_k$, то прямая сумма G этих групп также будет конечной группой и ее порядок n равен произведению порядков прямых слагаемых,

$$n = n_1 n_2 \ldots n_k. \qquad (16)$$

Действительно, число различных систем вида (13), у которых элемент a_1 может принимать n_1 различных значений, элемент a_2 принимает n_2 различных значений и т. д., определяется равенством (16).

Рассмотрим некоторые примеры.

Если порядок n конечной циклической группы $\{a\}$ разлагается в произведение двух взаимно простых натуральных чисел,

$$n = st, \quad (s, t) = 1,$$

то группа $\{a\}$ разлагается в прямую сумму двух циклических групп, имеющих соответственно порядки s и t.

Будем употреблять для группы $\{a\}$ аддитивную запись. Если положим $b = ta$, то

$$sb = (st)\,a = na = 0,$$

но для $0 < k < s$

$$kb = (kt)\,a \neq 0,$$

т. е. циклическая подгруппа $\{b\}$ имеет порядок s. Аналогично циклическая подгруппа $\{c\}$ элемента $c = sa$ имеет порядок t. Пересечение $\{b\} \cap \{c\}$ содержит только нуль, так как если $kb = lc$ при $0 < k < s$, $0 < l < t$, то

$$(kt)\,a = (ls)\,a,$$

откуда, так как числа kt и ls меньше n,

$$kt = ls,$$

что невозможно ввиду взаимной простоты чисел s и t. Наконец, существуют такие числа u и v, что

$$su + tv = 1,$$

а поэтому

$$a = v\,(ta) + u\,(sa) = vb + uc,$$

и, следовательно, любой элемент группы $\{a\}$ можно представить как сумму элементов из подгрупп $\{b\}$ и $\{c\}$.

Назовем абелеву группу G *неразложимой*, если ее нельзя разложить в прямую сумму двух или нескольких ее подгрупп, отличных от нулевой подгруппы. Конечная циклическая группа, порядок

которой является некоторой степенью простого числа p, называется *примарной* циклической группой, относящейся к простому числу p. Применяя несколько раз доказанное выше утверждение, мы получим, что *всякая конечная циклическая группа разлагается в прямую сумму примарных циклических групп, относящихся к различным простым числам.* Точнее, *циклическая группа порядка*

$$n = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s},$$

где $p_1, p_2, \ldots, p_s$ — *различные простые числа, разлагается в прямую сумму* s *циклических групп, имеющих соответственно порядки* $p_1^{k_1}, p_2^{k_2}, \ldots, p_s^{k_s}$.

Всякая примарная циклическая группа неразложима.

В самом деле, пусть дана конечная циклическая группа $\{a\}$ порядка p^k, где p — простое число. Если бы эта группа была разложимой, то, по (7), она обладала бы ненулевыми подгруппами, пересечение которых равно нулю. В действительности, однако, всякая ненулевая подгруппа нашей группы содержит отличный от нуля элемент

$$b = p^{k-1} a.$$

Для доказательства возьмем произвольный ненулевой элемент x нашей группы,

$$x = sa, \quad 0 < s < p^k.$$

Число s можно записать в виде

$$s = p^l s', \quad 0 \leqslant l < k,$$

где число s' уже не делится на p и, следовательно, взаимно просто с ним, а поэтому существуют такие числа u и v, что

$$s'u + pv = 1.$$

Тогда

$$(p^{k-l-1}u)\,x = (p^{k-l-1}us)\,a = (p^{k-1}us')\,a =$$
$$= p^{k-1}(1 - pv)\,a = (p^{k-1} - p^k v)\,a = p^{k-1}a - v\,(p^k a) = p^{k-1}a = b,$$

т. е. элемент b входит в циклическую подгруппу $\{x\}$.

Аддитивная группа целых чисел (т. е. бесконечная циклическая группа), а также аддитивная группа всех рациональных чисел являются неразложимыми группами.

Неразложимость обеих указанных групп вытекает из того, что в каждой из этих групп для любых двух ненулевых элементов существует ненулевое общее кратное, т. е. любые две ненулевые циклические подгруппы обладают ненулевым пересечением.

Заметим, что если операция в абелевой группе G называется умножением, то следует говорить не о прямой сумме, а о *прямом произведении*.

Мультипликативная группа отличных от нуля действительных чисел разлагается в прямое произведение мультипликативной группы положительных действительных чисел и группы по умножению, составленной из чисел 1 *и* -1.

Действительно, в пересечении указанных двух подгрупп нашей группы содержится лишь число 1 — единичный элемент этой группы. С другой стороны, всякое положительное число является произведением самого себя на число 1, всякое отрицательное число — произведением своей абсолютной величины на число -1.

§ 67. Конечные абелевы группы

Если мы возьмем любой конечный набор примарных циклических групп, некоторые из которых могут относиться к одному и тому же простому числу или даже иметь один и тот же порядок, т. е. быть изоморфными, то прямая сумма этих групп будет конечной абелевой группой. Оказывается, что этим исчерпываются все конечные абелевы группы:

Основная теорема о конечных абелевых группах. *Всякая конечная абелева группа G, не являющаяся нулевой группой, разлагается в прямую сумму примарных циклических подгрупп.*

Доказательство этой теоремы начнем с замечания, что *в группе G непременно найдутся ненулевые элементы, порядки которых являются степенями простых чисел.* Действительно, если некоторый ненулевой элемент x группы G имеет порядок l, $lx=0$, и если p^k, $k>0$, есть такая степень простого числа p, на которую число l делится,

$$l=p^k m,$$

то элемент mx отличен от нуля и имеет порядок p^k.

Пусть

$$p_1, p_2, \ldots, p_s \tag{1}$$

будут все различные простые числа, некоторые степени которых служат порядками некоторых элементов группы G. Обозначим через p любое из этих чисел, а через P совокупность элементов группы G, имеющих своими порядками степени числа p.

Множество P является подгруппой группы G. Действительно, в P входит элемент 0, так как его порядок есть $1=p^0$. Далее, если $p^k x=0$, то и $p^k(-x)=0$. Наконец, если $p^k x=0$, $p^l y=0$ и если, например, $k \geqslant l$, то

$$p^k(x+y)=0,$$

т. е. порядком элемента $x+y$ служит или число p^k, или делитель этого числа, т. е. во всяком случае некоторая степень числа p.

Беря в качестве p поочередно каждое из чисел (1), мы получим s ненулевых подгрупп

$$P_1, P_2, \ldots, P_s. \tag{2}$$

Группа G является прямой суммой этих подгрупп,

$$G = P_1 + P_2 + \ldots + P_s. \tag{3}$$

Действительно, если x — произвольный элемент группы G, то его порядок l может делиться лишь на некоторые простые числа из системы (1),

$$l = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s},$$

где $k_i \geqslant 0$, $i = 1, 2, \ldots, s$. Поэтому, как показано в конце предшествующего параграфа, циклическая подгруппа $\{x\}$ разлагается в прямую сумму примарных циклических подгрупп, имеющих соответственно порядки $p_1^{k_1}, p_2^{k_2}, \ldots, p_s^{k_s}$. Эти примарные циклические подгруппы лежат в соответственных подгруппах (2), и, следовательно, элемент x представляется в виде суммы элементов, взятых по одному во всех или некоторых из подгрупп (2). Этим доказано равенство

$$G = \{P_1, P_2, \ldots, P_s\},$$

аналогичное равенству (6) из предшествующего параграфа.

Для доказательства равенства, аналогичного равенству (7) из того же параграфа, возьмем любое i, $2 \leqslant i \leqslant s$. Тогда любой элемент y из подгруппы $\{P_1, P_2, \ldots, P_{i-1}\}$ имеет вид

$$y = a_1 + a_2 + \ldots + a_{i-1},$$

где элемент a_j, $j = 1, 2, \ldots, i-1$, лежит в подгруппе P_j, т. е. имеет порядок $p_j^{k_j}$. Тогда

$$(p_1^{k_1} p_2^{k_2} \ldots p_{i-1}^{k_{i-1}})\, y = 0,$$

т. е. порядком элемента y служит некоторый делитель числа $p_1^{k_1} p_2^{k_2} \ldots p_{i-1}^{k_{i-1}}$ и, следовательно, элемент y, если он отличен от нуля, не может содержаться в подгруппе P_i. Этим доказано, что

$$\{P_1, P_2, \ldots, P_{i-1}\} \cap P_i = 0,$$

что и требовалось доказать.

Заметим, что абелева группа, порядки всех элементов которой являются степенями одного и того же простого числа p, называется *примарной* относительно числа p. Примарные циклические группы являются частным случаем примарных групп. Таким образом, подгруппы (2) примарны. Они называются *примарными компонентами* группы G, а прямое разложение (3) — *разложением этой группы в примарные компоненты.* Так как подгруппы (2) определены

в группе G однозначным образом, то и *разложение группы G в примарные компоненты определено однозначно.*

Разложимость всякой конечной абелевой группы в прямую сумму примарных групп сводит, понятно, доказательство основной теоремы на случай конечной примарной абелевой группы P, относящейся к некоторому простому числу p. Рассмотрим этот случай.

Пусть a_1 будет один из элементов группы P, имеющих в ней наивысший порядок. Если, далее, в группе P имеются ненулевые элементы, циклические подгруппы которых пересекаются с циклической подгруппой $\{a_1\}$ лишь по нулю, то через a_2 обозначим один из элементов наивысшего порядка среди элементов с этим свойством; таким образом

$$\{a_1\} \cap \{a_2\} = 0.$$

Пусть уже выбраны элементы $a_1, a_2, \ldots, a_{i-1}$. Подгруппу группы P, порожденную их циклическими подгруппами, обозначим через $\{a_1, a_2, \ldots, a_{i-1}\}$,

$$\{\{a_1\}, \{a_2\}, \ldots, \{a_{i-1}\}\} = \{a_1, a_2, \ldots, a_{i-1}\}. \tag{4}$$

Она состоит, очевидно, из всех элементов группы P, которые могут быть записаны в виде суммы элементов, кратных элементам a_1, $a_2, \ldots, a_{i-1}$; будем говорить, что эта подгруппа *порождается* элементами $a_1, a_2, \ldots, a_{i-1}$. Обозначим теперь через a_i один из элементов наивысшего порядка среди тех элементов группы P, циклические подгруппы которых имеют равное нулю пересечение с подгруппой $\{a_1, a_2, \ldots, a_{i-1}\}$; таким образом

$$\{a_1, a_2, \ldots, a_{i-1}\} \cap \{a_i\} = 0. \tag{5}$$

Ввиду конечности группы P этот процесс должен остановиться; пусть это произойдет после того, как будут выбраны элементы a_1, $a_2, \ldots, a_s$. Если через P' мы обозначим подгруппу, порожденную этими элементами,

$$P' = \{a_1, a_2, \ldots, a_s\},$$

т. е.

$$P' = \{\{a_1\}, \{a_2\}, \ldots, \{a_s\}\}, \tag{6}$$

то, следовательно, циклическая подгруппа любого ненулевого элемента группы P имеет с подгруппой P' ненулевое пересечение.

Равенство (6) и равенство (5), справедливое для $i = 2, 3, \ldots, s$, показывают, ввиду (4), что подгруппа P' является прямой суммой циклических подгрупп $\{a_1\}, \{a_2\}, \ldots, \{a_s\}$,

$$P' = \{a_1\} + \{a_2\} + \ldots + \{a_s\}. \tag{7}$$

Остается доказать, что подгруппа P' на самом деле совпадает со всей группой P.

Пусть x — любой элемент группы P, имеющий порядок p. Так как

$$P' \cap \{x\} \neq 0,$$

а подгруппа $\{x\}$ не имеет ненулевых подгрупп, отличных от нее самой — напомним, что порядок подгруппы является делителем порядка группы, а число p простое, — то в действительности подгруппа $\{x\}$ содержится в подгруппе P' и, следовательно, x принадлежит к P'. Таким образом, все элементы порядка p из группы P входят в подгруппу P'.

Пусть уже доказано, что в подгруппу P' входят все элементы группы P, порядок которых не превосходит числа p^{k-1}, и пусть x — любой элемент из P, имеющий порядок p^k. Как показывает выбор элементов $a_1, a_2, \ldots, a_s$, порядки их идут не возрастая и поэтому можно указать такое i, $1 \leqslant i-1 \leqslant s$, что порядки элементов $a_1, a_2, \ldots, a_{i-1}$ больше или равны p^k, а при $i-1 < s$ порядок элемента a_i строго меньше этого числа, т. е. меньше порядка элемента x. Отсюда следует, ввиду условий, которым подчинен выбор элемента a_i, что если

$$Q = \{a_1, a_2, \ldots, a_{i-1}\},$$

то

$$Q \cap \{x\} \neq 0.$$

В предшествующем параграфе было доказано, однако, что всякая ненулевая подгруппа примарной циклической группы $\{x\}$ порядка p^k содержит элемент

$$y = p^{k-1} x. \tag{8}$$

Элемент y входит, следовательно, в пересечение $Q \cap \{x\}$, а поэтому и в подгруппу Q. Это позволяет записать y в виде суммы элементов, кратных элементам $a_1, a_2, \ldots, a_{i-1}$,

$$y = l_1 a_1 + l_2 a_2 + \ldots + l_{i-1} a_{i-1}. \tag{9}$$

Из (8) следует, что элемент y имеет порядок p. Поэтому

$$(pl_1) a_1 + (pl_2) a_2 + \ldots + (pl_{i-1}) a_{i-1} = 0,$$

т. е., ввиду существования прямого разложения (7),

$$(pl_j) a_j = 0, \qquad j = 1, 2, \ldots, i-1.$$

Число pl_j должно, следовательно, делиться на порядок элемента a_j, а поэтому и на число p^k, откуда вытекает, что l_j делится на p^{k-1},

$$l_j = p^{k-1} m_j, \qquad j = 1, 2, \ldots, i-1. \tag{10}$$

Пусть

$$z = m_1 a_1 + m_2 a_2 + \ldots + m_{i-1} a_{i-1}.$$

Это будет элемент из подгруппы Q, а поэтому и из подгруппы P', причем, ввиду (9) и (10),

$$y = p^{k-1} z. \tag{11}$$

Из (8) и (11) вытекает равенство

$$p^{k-1}(x-z)=0,$$

т. е. порядок элемента

$$t=x-z$$

не больше p^{k-1} и, следовательно, в силу индуктивного предположения t содержится в подгруппе P'. Поэтому и элемент x, как сумма двух элементов из P', $x=z+t$, принадлежит к подгруппе P'. Этим доказано, что все элементы порядка p^k из группы P содержатся в P'.

Наше индуктивное доказательство позволяет утверждать, следовательно, что все элементы группы P входят в подгруппу P', т. е. $P'=P$. Доказательство основной теоремы закончено.

В качестве побочного продукта мы получаем, что *конечная абелева группа тогда и только тогда будет примарной относительно простого числа p, если ее порядок является степенью этого числа p.* В самом деле, было показано, что всякая конечная примарная (по p) абелева группа P разлагается в прямую сумму примарных (по p) циклических групп, а поэтому порядок группы P равен произведению порядков этих циклических групп, т. е. является степенью числа p. Обратно, если конечная абелева группа имеет порядок p^k, где p — простое число, то порядок любого ее элемента будет делителем этого числа, т. е. также некоторой степенью числа p, а поэтому группа оказывается примарной относительно p.

Основная теорема еще не исчерпывает вопроса о полном описании конечных абелевых групп, так как пока не исключена возможность того, что прямые суммы двух различных наборов циклических групп, примарных по некоторым простым числам, могут оказаться изоморфными группами. На самом деле это не имеет места, как показывает следующая теорема:

Если конечная абелева группа G разложена двумя способами в прямую сумму примарных циклических подгрупп,

$$G=\{a_1\}+\{a_2\}+\ldots+\{a_s\}=\{b_1\}+\{b_2\}+\ldots+\{b_t\}, \qquad (12)$$

то оба прямых разложения обладают одним и тем же числом прямых слагаемых, $s=t$, и между прямыми слагаемыми этих разложений можно установить такое взаимно однозначное соответствие, что соответствующие слагаемые являются циклическими группами одного и того же порядка, т. е. изоморфны.

Заметим сначала, что если мы в первом, например, из прямых разложений (12) соберем прямые слагаемые, относящиеся к данному простому числу p, то их прямая сумма будет примарной (по p) подгруппой группы G и даже примарной компонентой этой группы, так как ее порядок равен наивысшей степени числа p, на которую делится порядок группы G. Объединяя этим способом прямые слагаемые в каждом из разложений (12), мы в обоих случаях получим

разложение группы G в примарные компоненты, единственность которого уже была отмечена выше.

Это позволяет доказывать нашу теорему в предположении, что группа G сама является примарной относительно простого числа p. Пусть нумерация прямых слагаемых в каждом из разложений (12) выбрана так, что порядки этих слагаемых идут не возрастая, т. е. элементы $a_1, a_2, \ldots, a_s$ имеют соответственно порядки

$$p^{k_1}, p^{k_2}, \ldots, p^{k_s},$$

причем

$$k_1 \geqslant k_2 \geqslant \ldots \geqslant k_s,$$

а элементы $b_1, b_2, \ldots, b_t$ — порядки

$$p^{l_1}, p^{l_2}, \ldots, p^{l_t},$$

причем

$$l_1 \geqslant l_2 \geqslant \ldots \geqslant l_t.$$

Если бы утверждение нашей теоремы не имело места, то нашлось бы такое i, $i \geqslant 1$, что

$$k_1 = l_1, \ \ldots, \ k_{i-1} = l_{i-1}, \tag{13}$$

но

$$k_i \neq l_i.$$

Понятно, что $i \leqslant \min(s, t)$, так как для каждого из разложений (12) произведение порядков всех прямых слагаемых равно порядку группы G. Покажем, что наше предположение приводит к противоречию.

Пусть, например,

$$k_i < l_i. \tag{14}$$

Обозначим через H совокупность элементов группы G, порядки которых не превосходят p^{k_i}. Это будет подгруппа группы G, так как если x и y — элементы из H, то и $x+y$, и $-x$ имеют порядки, не превосходящие числа p^{k_i}.

Заметим, что к подгруппе H принадлежат, в частности, следующие элементы:

$$p^{k_1-k_i}a_1, \ p^{k_2-k_i}a_2, \ \ldots, \ p^{k_{i-1}-k_i}a_{i-1}, \ a_i, \ a_{i+1}, \ \ldots, \ a_s.$$

С другой стороны, если $1 \leqslant j \leqslant i-1$, то элемент $p^{k_j-k_i-1}a_j$ имеет порядок p^{k_i+1} и поэтому в H не входит. Отсюда следует, что смежный класс $a_j + H$ (напоминаем, что мы используем аддитивную запись!) имеет, как элемент фактор-группы G/H, порядок $p^{k_j-k_i}$; таков же порядок его циклической подгруппы $\{a_j + H\}$. Докажем, что группа

G/H является прямой суммой циклических подгрупп $\{a_j+H\}$, $j=1, 2, \ldots, i-1$,

$$G/H=\{a_1+H\}+\{a_2+H\}+\ldots+\{a_{i-1}+H\}, \tag{15}$$

и поэтому ее порядок равен числу

$$p^{(k_1-k_i)+(k_2-k_i)+\ldots+(k_{i-1}-k_i)}. \tag{16}$$

Если x — произвольный элемент группы G, то существует запись

$$x=m_1a_1+m_2a_2+\ldots+m_sa_s.$$

Пусть для $j=1, 2, \ldots, i-1$

$$m_j=p^{k_j-k_i}q_j+n_j,$$

где

$$0\leqslant n_j<p^{k_j-k_i}. \tag{17}$$

Тогда

$$m_ja_j=q_j(p^{k_j-k_i}a_j)+n_ja_j,$$

а так как первое слагаемое правой части содержится в H, то

$$m_ja_j+H=n_ja_j+H.$$

С другой стороны,

$$m_ia_i+H=H, \quad \ldots, \quad m_sa_s+H=H.$$

Поэтому

$$\begin{aligned} x+H&=(m_1a_1+H)+(m_2a_2+H)+\ldots+(m_sa_s+H)=\\ &=(n_1a_1+H)+(n_2a_2+H)+\ldots+(n_{i-1}a_{i-1}+H). \end{aligned} \tag{18}$$

Пусть существует еще одна такая запись,

$$x+H=(n'_1a_1+H)+(n'_2a_2+H)+\ldots+(n'_{i-1}a_{i-1}+H), \tag{19}$$

где

$$0\leqslant n'_j<p^{k_j-k_i}, \quad j=1, 2, \ldots, i-1. \tag{20}$$

Тогда элементы

$$n_1a_1+n_2a_2+\ldots+n_{i-1}a_{i-1}$$

и

$$n'_1a_1+n'_2a_2+\ldots+n'_{i-1}a_{i-1}$$

лежат в одном смежном классе по H, т. е. их разность принадлежит к H и поэтому

$$p^{k_i}\left[(n_1-n'_1)a_1+(n_2-n'_2)a_2+\ldots+(n_{i-1}-n'_{i-1})a_{i-1}\right]=0.$$

Отсюда следует (так как первое из разложений (12) — прямое), что

$$p^{k_i}(n_j-n'_j)a_j=0, \quad j=1, 2, \ldots, i-1,$$

а поэтому число $p^{k_i}(n_j - n'_j)$ должно делиться на порядок p^{k_j} элемента a_j и, следовательно, разность $n_j - n'_j$ делится на число $p^{k_j - k_i}$. Отсюда, ввиду (17) и (20), следует, что

$$n_j = n'_j, \quad j = 1, 2, \ldots, i-1,$$

т. е. записи (18) и (19) тождественны. Этим доказано существование прямого разложения (15).

Аналогичные рассмотрения, проведенные для второго из прямых разложений (12), покажут, что эта же фактор-группа G/H обладает прямым разложением

$$G/H = \{b_1 + H\} + \{b_2 + H\} + \ldots + \{b_{i-1} + H\} + \{b_i + H\} + \ldots,$$

т. е., ввиду (13) и (14), ее порядок должен быть строго больше числа (16). Это противоречие доказывает теорему.

Полное обозрение конечных абелевых групп нами теперь уже получено. Именно, *берем всевозможные конечные наборы натуральных чисел*

$$(n_1, n_2, \ldots, n_k),$$

отличных от единицы, но не обязательно различных, причем каждое из этих чисел должно быть степенью некоторого простого числа. Каждому такому набору ставим в соответствие прямую сумму циклических групп, порядками которых служат числа из этого набора. Все полученные этим путем конечные абелевы группы будут попарно неизоморфными, а любая другая конечная абелева группа изоморфна одной из этих групп.

УКАЗАТЕЛЬ ЛИТЕРАТУРЫ

В указателе приведены книги по различным разделам алгебры, вышедшие на русском языке за последние тридцать пять лет. Некоторые из этих книг являются учебниками или учебными пособиями по университетским или педвузовским алгебраическим курсам, а другие — свободными сочинениями или же монографиями по отдельным вопросам, рассчитанными на хорошо подготовленного читателя.

Высшая алгебра

Сушкевич А. К., Основы высшей алгебры, изд. 4, Гостехиздат, 1941.

Окунев Л. Я., Высшая алгебра, изд. 2, «Просвещение», 1966.

Шапиро Г. М., Высшая алгебра, изд. 4, Учпедгиз, 1938.

Ляпин Е. С., Курс высшей алгебры, изд. 2, Учпедгиз, 1955.

Фаддеев Д. К. и Соминский И. С., Сборник задач по высшей алгебре, изд. 8, Физматгиз, 1964.

Виноградов С. П., Основания теории детерминантов, изд. 4, ОНТИ, 1935.

Линейная алгебра

Гельфанд И. М., Лекции по линейной алгебре, изд. 3, «Наука», 1966.

Мальцев А. И., Основы линейной алгебры, изд. 2, Гостехиздат, 1956.

Шилов Г. Е., Введение в теорию линейных пространств, изд. 2, Гостехиздат, 1956.

Проскуряков И. В., Сборник задач по линейной алгебре, изд. 3, «Наука», 1967.

Гантмахер Ф. Р., Теория матриц, изд. 3, «Наука», 1967.

Бохер М., Введение в высшую алгебру, ОНТИ, 1934.

Шрейер О. и Шпернер Е., Введение в линейную алгебру в геометрическом изложении, т. I, ОНТИ, 1934.

Шрейер О. и Шпернер Е., Теория матриц, ОНТИ, 1936.

Фаддеев Д. К. и Фаддеева В. Н., Вычислительные методы линейной алгебры, Физматгиз, 1960.

Фрезер Р., Дункан В. и Коллар А., Теория матриц и ее приложения к дифференциальным уравнениям и динамике, ИЛ, 1950.

Гуревич Г. Б., Основы теории алгебраических инвариантов, Гостехиздат, 1948.

Теория групп, колец и структур

Ван-дер-Варден Б. Л., Современная алгебра, ч. 1 и 2, Гостехиздат, 1947.

Шмидт О. Ю., Абстрактная теория групп, изд. 2, ОНТИ, 1933. (См. также Шмидт О. Ю., Избранные труды, Математика, Изд. АН СССР, 1959.)

Курош А. Г., Теория групп, изд. 3, «Наука», 1967.

Александров П. С., Введение в теорию групп, изд. 2, Учпедгиз, 1951.

Джекобсон Н., Теория колец, ИЛ, 1947.
Чеботарев Н. Г., Введение в теорию алгебр, Гостехиздат, 1949.
Биркгоф Г., Теория структур, ИЛ, 1952.
Сушкевич А. К., Теория обобщенных групп, ГНТИ Украины, 1937.
Окунев Л. Я., Основы современной алгебры, Учпедгиз, 1941.
Бэр Р., Линейная алгебра и проективная геометрия, ИЛ, 1955.
Ляпин Е. С., Полугруппы, Физматгиз, 1960.
Картан А. и Эйленберг С., Гомологическая алгебра, ИЛ, 1960.
Джекобсон Н., Строение колец, ИЛ, 1961.
Курош А. Г., Лекции по общей алгебре, Физматгиз, 1962.

Теория полей

Чеботарев Н. Г., Основы теории Галуа, ч. 1, ОНТИ, 1934.
Чеботарев Н. Г., Теория Галуа, ОНТИ, 1936.
Гекке Э., Лекции по теории алгебраических чисел, Гостехиздат, 1940.
Вейль Г., Алгебраическая теория чисел, ИЛ, 1947.
Чеботарев Н. Г., Теория алгебраических функций, Гостехиздат, 1948.
Ходж В. и Пидо Д., Методы алгебраической геометрии, тт. 1 и 2, ИЛ, 1954; т. 3, ИЛ, 1955.
Граве Д. А., Трактат по алгебраическому анализу, тт. 1 и 2, Изд. АН УССР, 1938—1939.

Непрерывные группы

Понтрягин Л. С., Непрерывные группы, изд. 2, Гостехиздат, 1954.
Чеботарев Н. Г., Теория групп Ли, Гостехиздат, 1940.
Шевалле К., Теория групп Ли, ч. 1, ИЛ, 1948, ч. 2, 3, ИЛ, 1958.
Вейль Г., Классические группы, их инварианты и представления, ИЛ, 1947.
Мурнаган Ф. Д., Теория представлений групп, ИЛ, 1950.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ